DOCUMENT RESUME

ED 350 350 TM 019 128

AUTHOR Brown, William L.
TITLE Analysis of Spring 1990 Field Tryouts of MEAP Essential Skills Mathematics Test Items Using a Partial Credit Model.
PUB DATE Apr 92
NOTE 71p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Ability; Computer Assisted Testing; *Cutting Scores; Error of Measurement; Estimation (Mathematics); *Field Studies; Goodness of Fit; Grade 10; High Schools; *High School Students; Item Response Theory; Mathematical Models; *Mathematics Tests; Maximum Likelihood Statistics; *Test Items
IDENTIFIERS Dichotomous Variables; *Essential Skills Mathematics Tests; Michigan Educational Assessment Program; One Parameter Model; *Partial Credit Model; Rasch Model

ABSTRACT

The partial credit model of G. N. Masters (1982), a one-parameter unidimensional polychotomous Rasch model, was used to reduce the error of measurement, particularly for students near the cut score, and to permit measurement to reflect the actual ability of a student more accurately by reducing the degree of misfit for students near the cut scores. Two implementations of the extended Rasch model, the FACETS and BIGSCALE programs, were used in the application of polychotomous scoring to the Michigan Educational Assessment Program (MEAP) spring 1990 item tryouts of the Essential Skills Mathematics Tests. FACETS produces a Rasch analysis of many-faceted data through iterative maximum likelihood estimation. The dichotomous analysis was made with BIGSCALE, which fits the one-parameter Rasch measurement model to the data. Item responses were examined for approximately 264 10th-grade students. The analysis did not show a major difference in the students exceeding the cut score under the partial credit model as compared with dichotomous scoring, but it did show that error of measurement was reduced near the cut score. The degree of misfit to the model was slightly increased for students near the cut score. Appendix A contains portions of the examination. Appendixes B, C, D, and E give analysis results. Thirty-eight graphs and 47 references are included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
William L. Brown
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED350350

Analysis of Spring 1990 Field Tryouts of MEAP Essential Skills Mathematics Test Items Using a Partial Credit Model

William L. Brown

Michigan State University

Presented to the
American Educational Research Association
San Francisco, California
April 21, 1992

Running Head: Partial Credit Model

TM019278

BEST COPY AVAILABLE

Introduction

The objectives of this study were to use Master's (1982) Partial Credit Model (PCM), a one-parameter Unidimensional Polychotomous Rasch Model (UPRM), to (1) reduce the error of measurement, especially for students near the cut score, and (2) permit the measurement to more accurately reflect the actual ability of student (particularly at scores at or below the "guessing" range) by reducing the degree of misfit from the UPRM for students near the cut-score.

In a conventional multiple-choice test item, there is one correct response and three or four incorrect responses, establishing a dichotomy in which some examinees are declared to have "passed" the item and all others are judged to have "failed" it. In developing test items, the incorrect response choices (foils) are often designed to represent different types or levels of errors which will lead a novice to conclude that one of the defective responses is actually the correct one. Since there is almost never a situation in which an examinee is either completely knowledgeable or completely incompetent, it is theoretically possible to design foils which form an ordinal scale, with each foil reflecting a degree of mastery relative to the attribute being tested.

By estimating the approximate competency level associated with each foil, a test can produce more information than one in which each answer is scored dichotomously as either "correct" or "incorrect." It has been recognized that this has particularly important implications in making pass/fail decisions in criterion- referenced measurement. The added information also has diagnostic value by enabling the measurement of partial mastery. Test scoring schemes which give credit for "partially correct" responses fit into many classifications; this study is based upon Master's (1982) Partial Credit Model (PCM), a one-parameter Unidimensional Polychotomous Rasch Model (UPRM).

BACKGROUND

Over the past several years, traditional multiple-choice tests have been criticized for being narrow in scope and for providing information only at the knowledge and/or recall levels and giving little or no information about the

thought processes of the individual being tested (Stiggins & Bridgeford, 1985; Resnick & Resnick, in press). Shepard (1991) has noted that:

> Under pressure to raise test scores, the known limitations of multiple-choice tests have become greatly exaggerated. They become less valid indicators of what students know (because scores can go up without a commensurate gain in achievement); and more seriously, when multiple-choice tests become the focus of instructional effort, they have a negative effect on teaching and learning. (p. 21)

Despite these criticisms, multiple-choice tests continue to be an indispensable part of the assessment program in the schools. Since multiple-choice tests are easily scored by computer, they are efficient to use in addition to offering high reliability and validity. Several researchers have used polychotomous scoring to obtain more information about an individual's ability, while retaining the benefits of multiple-choice testing (Smith, 1987; Bock, 1972; Hambleton, *et al*, 1970). These authors have shown that the degree of correctness of an answer can be quantified and used as an additional source of information about an individual's ability.

Many teacher-made, hand-scored examinations (particularly at the secondary school and college levels) are graded to award credit for partially correct responses. In particular, examinations in mathematics, science, and engineering are often so scored (Crittendon, 1984). Polychotomous scoring can (1) provide more evidence of an examinee's partial development of the underlying trait, (2) improve the accuracy of the measurement, and (3) provide as much information with fewer items than would be needed for comparable results from dichotomous scoring (Haladyna, 1990). Furthermore, the diagnostic utility of the information obtained from partially correct responses has been demonstrated in several studies (Levine & Drasgow, 1983; Thissen, 1976).

Despite growing disenchantment among some educators with machine-scored multiple-choice tests (Gifford, *et al*, 1990), the efficiency and objectivity of these instruments provide strong incentives for continuing their use. In a study designed to determine whether open-ended items provide more predictive validity than dichotomously-scored multiple-choice items, chemical engineering students at Purdue University were examined with both grader-

scored (open-ended) and machine-scored (multiple-choice) quizzes. It was found that (Kessler, 1988):

> ... grader-scored quizzes are not any better correlated (than machine-scored quizzes) with the usual grade-index criteria of student performance. ... Although there may be a difference in evaluations using quizzes designed for machine- and grader-scored testing, it appears that other variables are so much more important that they obscure any practical differences (p. 709).

In an effort to overcome the presumed limitations of dichotomously-scored multiple choice items, various methods of "differential option weighting" have been investigated (Yan, 1989). These include answer-until-correct, examinee's confidence weighting, distractor elimination, and response weighting. Each of these methods has been explored further using linear and non-linear models to determine the weights to be assigned to each option (Haladyna, 1988). Although there have been numerous techniques and formats developed to grant partial credit through various schemes, this paper is restricted to paper-and-pencil tests using standard multiple-choice items that have been administered under directions to "find the BEST answer." Logically, the correlation between each option choice and the total test score (the point-biserial coefficient) provides a measure of the degree of correctness of the option as viewed by those taking the test. One method for determining distractor weights, then, is to arrange the option choices in descending point-biserial order (cf. Haladyna, 1990, pp. 232f). A more comprehensive overview is included in Haladyna's (1988, 1990) papers.

Beyond the utility of partial credit scoring for instructional purposes, Haladyna (1990) has noted that:

> For any testing program intended for licensure, certification, competency, or proficiency, ... empirical option-weighting typically produced slightly more reliable domain score estimates and more consistent pass/fail decisions than number-correct scoring, particularly in the lower half of the test score distribution. (p. 231)

The 1979 Joint College Entrance Examination in Taiwan, Republic of China, used a "multiple answer" format in which there were several response options for each item, some of which were correct and some incorrect. To receive credit for an item, <u>all</u> response choices had to be marked correctly.

Data from administration of this test were later re-analyzed to give partial credit for students who misclassified one or more options on a given item (Hsu, 1984). The conclusion was that the tests were easier, more reliable, and more discriminating by granting partial credit, although the gains were small. The limited gains were attributed to "ceiling effects" because these tests were already very reliable and discriminating without the partial credit scoring.

Another example comes from a study of factors which influence the learning of classical mechanics . It was discovered that (Champagne *et al*, 1980):

> ... on the basis of common-sense world experiences, students often develop conceptions about the physical world which are strongly held and which interfere with the learning of new conceptual relationships during physics instruction. ... instruction in classical mechanics can be improved by continuously encouraging the students to reject an Aristotelian system of beliefs and to adopt a Newtonian paradigm. (pp. 1074 and 1078)

Since paradigm shifts are usually resisted or entered into reluctantly (Kuhn, 1962), diagnosis of student progress in learning classical mechanics is strengthened by designing test items which recognize partial learning of the theoretical concepts of Newton as contrasted with the "common-sense" notions of Aristotle which students acquire early in life through their personal observations.

In a report on developments in the area of computerized adaptive testing (CAT) using partial credit scoring, it was noted that (Dodd *et al*, 1989):

> ... research with the graded response model suggests that it may be possible to implement CAT successfully using item pools that are substantially smaller than the pools required for dichotomous items. The apparent reason is that polychotomous scoring of items provides more information across the full range of the Q scale, which reduces the possibility that gaps will occur in the pool. (p. 141)

The fallacy of combining unrelated attributes into a single result is one of the most severe criticisms of common measurement "tools" like performance ratings, which have been used for many years in personnel evaluations and

are currently widely touted as "authentic assessment." These instruments are prone to the flaw of "halo," in which the rater tends to appraise diverse characteristics as if they were common attributes. Supervisor's ratings tend to be unreliable because (Hunter & Schmidt, 1990):

> ... the idiosyncrasy of that supervisor's perceptions is a part of the error of measurement in the observed ratings. Extraneous factors that may influence human judgment include friendship, physical appearance, moral and/or life-style conventionality, and more (p. 65).

For example, police officers who are considered to be effective investigators may receive high ratings as report writers, while a less effective investigator who writes far superior reports (where the reports are judged anonymously) may be rated as a poor report writer. Few measurement instruments exist which can clearly measure and report *distinct* attributes of the subject being measured (Denny, 1990).

How does a test "earn the privilege" of being sound enough that its "scores" can be said to represent true "measurement?" *Unidimensionality* is clearly desirable if strong measurement conclusions are to be drawn about the scores obtained. Wright & Linacre (1989) have noted that:

> ... unidimensionality is a meaning of measurement. No actual test can be perfectly unidimensional. ... This reality is encountered by every science. Physicists' corrections for the unavoidable multidimensionalities they encounter are an integral part of their experimental technique. (p. 2)

To argue, however, that a measurement model is "robust to violations of unidimensionality" does not relieve one from responding substantively to questions about dimensionality. Wright & Linacre go on to state that:

> If a test containing a mixture of medical and law items is used to make a single pass-fail decision, then the examination board, however inadvertently, has decided to use the test as if it were unidimensional. ... Their unidimensional behavior does testify that they are making medicine and law exchangeable and hence identical there. (p. 3)

The issue of *unidimensionality* is critical to any discussion of measurement, but it is *especially* so in the case of the Partial Credit Model

(PCM). It would make very little sense, for instance, to give someone credit for partial understanding of medicine if one is using foils which represent partial knowledge of law. This should become clearer later in this presentation.

Methodology

Models

Rasch's dichotomous model for the probability of a correct response F_{ni} by person n (with ability b_n) pitted against an item i (with difficulty d_i) is:

$$F_{ni} = \exp(b_n - d_i)/(1 + \exp(b_n - d_i)) \quad (1)$$

An extension of simple right/wrong scoring consists of identifying one or more intermediate levels of performance on an item and awarding partial credit for reaching one of the intermediate levels (Wright & Masters, 1982, p. 40). The dichotomous Rasch model may be generalized to a model for partial credit by representing each pair of adjacent response opportunities in a string of ordered categories as a dichotomous element governed by $\exp(b_n - d_{ij})/(1 + \exp(b_n - d_{ij}))$, where j indicates the step required in item i for person n to advance from the lower of the adjacent pair to the higher.

Thus, instead of a single equation defining the probability of a correct response, we have a family of equations representing the probability of an examinee selecting a given step of a "scale" of response choices. The difficulty, d_{ij}, of the jth step of the ith item can be written as

$$d_{ij} = d_i + t_{ij}, \quad (2)$$

where d_i is the base difficulty of item i and t_{ij} is the incremental difficulty (from the base) of the jth step. Therefore, the probability of a correct response is:

$$F_{nij} = \exp(b_n - (d_i + t_{ij}))/(1 + \exp(b_n - (d_i + t_{ij})) \quad j = 1, 2, \ldots, m_i \quad (3)$$

Instrumentation

Two implementations of the extended Rasch models were used in the

application of polychotomous scoring to the MEAP 1990 Spring mathematics item tryouts. They are FACETS (Wright & Linacre, 1990) and BIGSCALE (Wright & Linacre, 1989).

FACETS is a computer program which produces a Rasch analysis of many-faceted data through iterative maximum-likelihood estimations based upon models such as

$$\log () = Bn - (Di + Tj) \qquad (4)$$

where Bn is the ability of student n, Di is the difficulty of item i, and Tj is the step difficulty of response choice j. In quantitative terms, the probability of choosing the *k*th response category is

$$F_{nik} = \exp \sum_{j=0}^{k} (b_n - (d_i + t_{ij})) \Big/ \sum_{k=0}^{m} \exp \sum_{j=0}^{k} (b_n - (d_i + t_{ij})), \qquad (5)$$

where $t_{i0} \equiv 0$. FACETS was designed primarily for analyzing Likert-type rating scale data. It provides estimates of a difficulty parameter and a set of "threshold" values (t_{ij}) independently for each item. In effect, the program treats the set of response choices for each item as a unique rating scale, in which the worst response is given a rating of "-1" and the best response a rating of "+1." The other choices are positioned, by FACETS convention, along the scale between -1 and +1 through an iterative joint maximum likelihood method (Linacre, 1989). Other scale choices (i.e., other than -1 to +1) are available, however, and should be considered in future research.

The dichotomous analysis was made using BIGSCALE, a program which fits the one-parameter Rasch measurement model to the data. Results from the FACETS analysis were used to enter data into BIGSCALE for partial credit scoring with uniform steps for all items (i.e., using weights of 0, 1, 2, and 3 for the least desirable foil to the correct answer response, respectively).

Approximate *t*-statistics are computed by both FACETS and BIGSCALE, which summarize an individual's or an item's fit to the UPRM. A large positive fit statistic suggests that the person ability or item difficulty should be inspected for unexpected failures or successes (Wright & Masters, 1982). A large negative fit statistic indicates a consistent Guttman-like response pattern; a person with

statistic indicates a consistent Guttman-like response pattern; a person with such a fit statistic will have (a) correct scores on those items for which his ability is greater than the item difficulty, (b) incorrect scores for those items whose difficulties are above his ability, and (c) partial credit scores when his ability and item difficulties are approximately equal (Hillocks & Ludlow, 1984). Large negative fit statistics, then, demonstrate that the rating-scale property of the test items has been confirmed.

Content Analysis

This study examined the item responses of a sample of approximately 264 tenth grade students who were part of the Spring 1990 item tryout population for Form 3A1 of the MEAP Essential Skills Mathematics Tests. This test form included 134 items. Analysis of these items using FACETS showed that only 16 items had foils which appeared to represent varying degrees of understanding of the underlying concepts. (cf. Appendix A for the test items used in this study.)

Three examples will be used to illustrate the types of items which appear to fit the partial credit model. The first example (cf. Figure 6) asks:

2. Which of these numbers has the GREATEST value?

A 6.575×10^{-3}

B 1.27×10^{-2}

C 5.40×10^{-4}

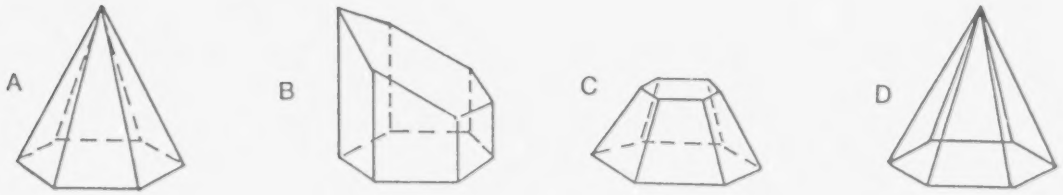
D 5.01×10^{-1}

"**D**" is the correct response, since 10^{-1} is the highest exponent and all of the numbers are stated in proper scientific notation (i.e., all of the numbers are stated in the form $\mathbf{X \times 10^{n}}$, where **X** is a decimal between 1 and 10 and **n** is the power of 10 which, when multiplied by **X**, gives the value of the number). **C** would attract those who do not know the meaning of the *minus sign* before the exponent, whereas **A** is designed to attract the student who does not

understand exponents and only looks at the **X** values in the response choices. **B** appears to be a "throw-away," since one would need to lack a fundamental understanding of the meaning of the word "GREATEST" to select this response. This item had a *p*-value of 0.50, and a point-biserial correlation of 0.49. **C** and **A** were each selected by about 21% of all the examinees, with the more able choosing **C** and the less able, **A**. Only 6% of all examinees chose **B**. Among the bottom 40% of the students, **A** was the favorite choice, while **C** was next highest.

Another example (Figure 8) asks:

45. Which of these is an appropriate drawing of a hexagonal pyramid?



For this example, **A** is the correct ("appropriate drawing") answer. **C** is close to being correct, but it is a *truncated* hexagonal pyramid. **D** could easily be chosen by an examinee who did not understand that lines which cannot be seen from the vantage point of the observer are shown as *dashes*. **B** is not even a pyramid. This item had a *p*-value of 0.41, and a point-biserial correlation of 0.49. **C** was selected by 25% of all examinees, **D** by 19%, and **B** by 10%. Among the bottom 40% of the examinees, **C** was the top choice, while **D** was favored by the middle 20% of the group.

The final example, in Figure 9, asks:

56. What is the probability of getting three heads in a row tossing a fair coin?

A 3/2 **B** 1/2 **C** 1/4 **D** 1/8

D is the correct answer. **B** is plausible to examinees who think only of the probability of getting *heads* on a single toss of a coin, while **A** indicates a moderate understanding of the problem but a misunderstanding of how to compute the joint probability (i.e., the probability of a 50-50 event occurring three times in a row, which is .50 * .50 * .50, not .50 + .50 + .50). This item had a *p*-value of 0.22 and a point-biserial correlation of 0.41. **B** was selected by 32% of all examinees, and was the favored choice of all but the top quintile (which chose **D**, the correct answer).

Data Analysis

Part (a) of Figures 4 through 21 are Item Characteristic Curves (ICCs) which illustrate how students with different raw scores on the 16-item testlet performed on the items which appeared to produce "partial credit" information. The numbers on each vertical axis represent the proportion of students in each respective quintile who selected each of the 4 answer choices. Part (b) of each figure is an "Item Response Map" which is an alternative method for presenting the data from Part (a) using the cumulative proportions responding to the items. The advantage of the Item Response Map is that it demonstrates whether or not the response categories actually form a "scale" of logistic Item Characteristic Curves, such as those illustrated in Figure 1(b). In each case, the numbers along the horizontal axis represent groups of students in the five quintiles of total raw scores. The correct option is shown in parentheses for each item.

The analysis of the test results indicated that these sixteen items produced moderately reliable results for a very difficult test, independent of scoring technique, with KR-20 being about .73 for all three analyses performed. Elaborating on this conclusion, the following subsections describe the results from dichotomous and polychotomous scoring of the sixteen-item "testlet." In order to provide a moderately stringent test of the PCM, a cut score of 25% correct was chosen for this study. This places some of the examinees who passed the test in the guessing range, since four options per item produces a guessing level of 25% correct. Since errors of measurement and model misfit most severely affect students in the low-score range of a test, this situation simulates Haladyna's (1990) conditions. For students near the cut score, comparisons are made of classification results, testlet reliability, and misfits to

the UPRM.

In order to compare results of dichotomous and PCM scoring, the cut score was set for all three analyses using equipercentile scaling. This produced a passing score of four items correct (out of 16 possible) on the dichotomously scored analysis, and 35 points (out of a possible 48) on the PCM-scored analyses.

BIGSCALE Dichotomous Scoring Analysis

BIGSCALE uses an iterative unconditional maximum likelihood procedure to determine the best-fit logistic curve for the test results. After a satisfactory fit is obtained, the program computes a modified *t*-statistic on each item response curve and each person's response pattern to determine how well each item and examinee fits the model. Since the test checks the fit of the item (or person) to the model, a two-tailed test is used (critical value is $t = 1.960$, $\alpha = .05$).

For the sixteen-item testlet, two items produced fit statistics which exceeded the critical value (cf. Appendix B, p. B-7). Items 62 (Figure 13, Area Problem) and 92 (Figure 16, Robot Selection) were very difficult items which elicited random-like responses from nearly all but the highest-scoring students.

Four persons (out of 262 tested) had fit statistics which exceeded the critical value (Appendix B, p. B-6). Students 18, 74, 110, and 131 had item response patterns which were `0000010000100011`, `0010110011001100`, `1000100010101011`, and `1000011011011100` (compared with an expected pattern such as `1111111100000000`, where the items are arranged from easiest to most difficult; 1 represents "correct," and 0 represents "incorrect"). These students missed at least 5 of the 8 easiest items, and got at least 4 of the 8 most difficult items correct. In contrast, student 217 had a fit statistic of 0 (response pattern = `1001111100001000`), and student 256 had a statistic of -2.2 (pattern = `1111111010000000`). The more negative the fit statistic, the closer the student's response pattern approximates a Guttman scale (i.e., a string of 1's followed by a string of 0's).

A large negative *t*-statistic does represent a misfit to the Rasch model, which could theoretically cause prediction errors for students whose ability levels are not close to the difficulty level of the item. Item 18 (Figure 5) is an

example of such an item with a large negative *t*-statistic and a difficulty calibration of 0.28 logits. Students with ability calibrations of .30 logits (the closest to the item calibration value) performed exactly as the model predicts on this item, with 54% getting the item right (N=13 students). Students with ability levels .7 logits above or below the item difficulty should score 75% correct and 25% correct, respectively; the actual figures are 75% correct (ability = .95 logits; N=20) and 0% correct (ability = -.31 logits; N=14). The model fit is still good for the more able students, but totally wrong for the less able examinees.

At the student level, the strong Guttman-like response of student 256 (pattern = 1111111010000000) is a good example. As a student encounters items on the test which most closely match his or her ability level, there is a 50-50 chance of a correct response. For a Guttman-like response, student 217 (with a fit-statistic of 0) would have been expected to get the first 7 items correct and the last 9 items incorrect. With the BIGSCALE analysis, the first item and the last three items received the expected responses (1 and 0, respectively); the middle 10 items form a Guttman scale; the intervening two items (second and thirteenth) received unexpected responses. Since the fit statistic is zero for this student, the response pattern reflects a very good fit to the Rasch model.

Altogether, 10.6% of the students in this sample exhibited a Guttman-scale response pattern with $t<-1$ (Appendix B, p. B-1). Another 10% had $t>1$, indicating a moderate lack of fit to the Rasch model, including 1.5% with fairly severe misfit ($t>1.96$). The remainder fit the model very well. Of the students at the 25th percentile (scoring 4 out of 16 items correct), the average *t*-statistic was +.48 with a mean-square error of .63 logits.

BIGSCALE Partial Credit Scoring Analysis

In the BIGSCALE Partial Credit Model analysis, there were four items (out of 16 tested) which had fit statistics which exceeded the critical value (cf. Appendix C, p. C-7). Item 91 (Figure 15, Hydrogen Peroxide) was a very difficult item which seemed to elicit almost random responses from all but the highest-scoring students, while items 10 (Figure 6, Income/Expense Trend) and 15 (Figure 7, Sum of -3, -15) were very easy. Item 62 (Figure 13, Area Problem) was of moderate difficulty.

Response patterns were generated for each student, with items

arranged in order of easiest to most difficult. Seven persons (out of 262 tested) had fit statistics which exceeded the critical value. These were students 58, 199, 5, 110, 108, 3, and 227, whose response patterns were 0320223002032213, 0322020300321223, 0002001130320022, 3011330120332323, 3302 00 00200303, 3121010330321332, and 3313330133002133, (compared with a Guttman-like 3332222211111000; **3** represents "fully correct", **2** "high-partially correct", **1** "low-partially correct", **0** represents "incorrect," and **blank** means the item was skipped). These students selected the "incorrect" (worst) response for at least 3 of the 8 easiest items, and got at least 2 of the 8 most difficult items "fully correct." In contrast, student 65 had a fit statistic of 0 (response pattern = 3133322320321223), and student 219 had a statistic of -2 (pattern = 3333111121221222). The more negative the fit statistic, the closer the student's response pattern comes to being a perfect Guttman scale (i.e., a string of 3's, followed by a string of 2's, followed by a string of 1's, followed by a string of 0's).

Altogether, 10.7% the students in this sample exhibited a Guttman-scale response pattern with $t<-1$ (Appendix C, p. C-1). Another 15.3% had $t>+1$, indicating a moderate lack of fit to the Rasch model, including 2.7% with a severe misfit, $t > 1.96$ (Appendix C, p. C-6). The remainder fit the model very well. Of the students scoring at the 25th percentile (25 points out of a possible 48), the average *t*-statistic was +.88 with a mean-square error of .24 logits. Note that the logit error (variability) decreased by 62% from the dichotomous value, while the *t*-statistic (goodness-of-fit) increased by 83%. Therefore, the measurement precision increased significantly, at the expense of decreased fit to the Rasch model.

FACETS Rating Scale Analysis

In the FACETS Rating Scale Model analysis, there was one item (out of 16 tested) with a fit statistic which exceeded the critical value (cf. Appendix D, p. D-7). Item 91 (Figure 15, Hydrogen Peroxide) was a very difficult item which severely misfitted the model for seven students (3% of the sample).

Twenty-eight persons (out of 262 tested) each had a fit statistic on one item which exceeded the critical value (Appendix D, p. D-6). Five students had high overall *t*-statistics indicating severe misfit to the model. These were among the same students who misfitted the BIGSCALE Partial Credit Model.

Altogether, 13.4% of the students in this sample exhibited a Guttman-scale response pattern with $t<-1$ (Appendix D, p. D-1). Another 15.3% had $t>+1$, indicating a moderate lack of fit to the Rasch model. (FACETS does not produce the same precision on the *t*-statistic as BIGSCALE produces.) Only 2% of the students had fit statistics which exceeded the critical value. The remainder fit the model very well. Of the students scoring at the cut-score of 25 out of 48 points, the average *t*-statistic was +.60 with a mean-square error of .24 logits. Note that the logit error was decreased by 62% just as it was with the BIGSCALE Partial Credit analysis, while the *t*-statistic was increased by 25%. Therefore, the precision of measurement was the same with FACETS and the BIGSCALE Partial Credit analysis, but the fit to the Rasch model is much better using FACETS.

Discussion

Results

The FACETS Rating Scale Analysis and the BIGSCALE dichotomous and PCM analyses of the 16-item testlet at the *item* level showed item-separation reliabilities of 0.96; at the *person* level, there was a person-separation reliability of .73 (cf. Appendix E).

Four students (1.6%) "passed" the test under the PCM who had "failed" under dichotomous scoring. Also, nine students (3.5%) "failed" the test under the PCM who had "passed" under dichotomous scoring. The average fit statistics for the newly passing students were reduced from -.35 to -.83, indicating better fit to the model. For the newly failing students, however, the average fit statistic *increased* from .51 to .88, meaning that the students who were "demoted" under PCM scoring experienced a greater degree of misfit than they did under dichotomous scoring. Overall, 5.1% of the students had their classification changed under PCM scoring and experienced an average increase of .10 in *t*-statistic (to an average *t* of +.13). Thus, the fit statistic for low-scoring students was not reduced using the PCM scoring, disconfirming the second hypothesis.

Overall, the root-mean-square error using dichotomous scoring was 0.70; using the BIGSCALE PCM scoring, it was reduced to 0.37, close to the value of 0.38 obtained using the FACETS Rating Scale Model. At the cut-score (the 25th percentile), the RMS error with dichotomous scoring was 0.63 logits; the

BIGSCALE PCM and Facets Rating Scale Analysis RMS error at the cut-score was 0.24 logits. This confirmed the first hypothesis (that the error of measurement at the cut-score was reduced using the Partial Credit Model).

Impact

Although the present analysis did not show a major difference in the students exceeding the cut-score under PCM as compared with dichotomous scoring (a difference of only 5.1%), it did show that the error of measurement was reduced near the cut-score, while the degree of misfit to the model was slightly increased for students near the cut-score. These results have potential practical significance for decisions regarding student performance.

Had the test been designed for PCM scoring, and had it not been quite so difficult (and embedded within a much larger test), it is possible that the results might have been more conclusive.

Limitations

The results of this study are limited because the test instrument which was used was not designed with the Partial Credit Model in mind. Preliminary analysis of the point-biserials indicated, however, that several of the items produced item characteristics similar to those needed for the partial credit model to work. About an eighth of the items in the test appeared to lend themselves to partial credit scoring. Even among cases where the items fit the partial credit model, the discrimination among the foils was too small to constitute different (and perhaps discrete) stages of learning.

Another limitation is the implied assumption that thinking processes are discrete and that there is a unique way to obtain the correct answer. Current developments in cognitive research indicate that there are multiple ways to approach a problem (Resnick & Resnick, in press), so this must be considered when developing foils for test items. Items which have non-unique approaches to their solutions may not lend themselves well to partial credit scoring.

In addition to these limitations, the results can be generalized only to a 10th-grade population such as that used in the MEAP 1990 spring item tryouts. Since the test was also used with 9th graders, it is possible that further analysis could determine that the findings apply to younger test-takers as well.

Recommendations

This study demonstrated that, for the data set used in the analysis, the error of measurement at the cut-score was reduced using the PCM, whereas the fit to the Rasch model was not improved with partial credit scoring. It is postulated that the fit to the PCM was not improved because the items were not designed with partial credit in mind. To eliminate this concern, the study should be extended using test items designed for PCM scoring. At the same time, that the items should be administered as a stand-alone unit, not as part of a larger test.

Since the PCM can be implemented easily using computer software, the quality of decisions could be improved relatively easily in large-scale assessments using such a scoring regimen. With the continuing growth of state

assessment programs, and the prospect for a national examination, the efficiency of multiple-choice testing and the information quality of PCM could be combined to provide useful diagnostic information. This information could be used on the large scale for curriculum planning, and at the individual student level for individual pupil diagnosis.

This study did not consider whether or not the foils actually represented ordinal scales of learning or performance, and no information was available on external measures of student performance. Estimates of test score validity need to be obtained, to see if the PCM provides improved measurement over that provided by similar dichotomously-scored instruments.

References

AERA, APA, & NCME (1985). *Standards for Educational and Psychological Testing.*

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43,* 561-573.

Bauer, E. L. (1989). Sense and nonsense about school quality indicators. Unpublished manuscript. Oakland Schools.

Ben-Shakhar, G., & Sinai, Y. (1991). Gender differences in multiple-choice tests: The role of differential guessing tendencies. *Journal of Educational Measurement, 28,* 23-35.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37,* 29-51.

Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education, 15,* 179-202.

Champagne, A.B., Klopfer, L.E., & Anderson, J.H. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics, 48,* 1074-1079.

Champagne, A.B., Klopfer, L.E., & Gunstone, R.F. (1982). Cognitive research and the design of science instruction. *Educational Psychologist, 17,* 31-53.

Choppin, B. H. L. (1985). Lessons for psychometrics from thermometry. *Evaluation in Education, 9,* 9-12.

Crittendon, J. B. (1984). Partial credit: Not a God-given right. *Engineering Education*, 288.

De Ayala, R.J., Dodd, B.G., & Koch, W.R. (1991). Partial Credit Analysis of Writing Ability. *Educational and Psychological Measurement, 51*, 103-114.

Denny, G. S. (April, 1990). A comparison of rater calibration methods. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Dodd, B.G., Koch, W.R., & De Ayala, R.J. (1989). Operational characteristics of adaptive testing procedures using the graded response model. *Applied Psychological Measurement, 13*, 129-143.

Donlon, T.F., & Fischer, F.E. (1968). An index of an individual's agreement with group-determined item difficulties. *Educational and Psychological Measurement, 28*, 105-113.

Duncan, O. D. (1984). **Notes on social measurement: Historical & critical**. NY: Russell Sage Foundation.

Ebel, R.L. (1968). Blind guessing on objective achievement tests. *Journal of Educational Measurement, 5*, 321-325..pa

Ferrara, S., & Walker-Bartnick, L. (1989), Constructing an essay prompt bank using the partial credit model. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, 1989.

Fisher, G.H. (1987). Applying the principles of specific objectivity and of generalizibality to the measurement of change. *Psychometrika, 52*, 565-587.

Gifford, B. R., Madaus, G. F., *et al* (1990). *From Gatekeeper to Gateway: Transforming Testing in America*. Chestnut Hill, MA: National Commission on Testing and Public Policy.

Haladyna, T. M. (April, 1988). Empirically based polychotomous scoring of multiple-choice items: A review. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, 1988.

Haladyna, T. M. (1990). Effects of empirical option weighting on estimating domain scores and making pass/fail decisions. *Applied Measurement in Education, 3*, 231-244.

Hambleton, R. K., Roberts, D. M., & Traub, R. E. (1970). A comparison of the reliability and validity of two methods of assessing partial knowledge on multiple choice tests. *Journal of Educational Measurement, 7*, 75-82.

Hillocks, G., & Ludlow, L.H. (1984). A taxonomy of skills in reading and interpreting fiction. *American Educational Research Journal, 21*, 7-24.

Hunter, J. E., & Schmidt, F. L. (1990). **Methods of Meta-Analysis: Correcting Error and Bias in Research Findings**. Newbury Park, CA: Sage Publications, Inc.

Hsu, T.C., Moss, P.A., & Khampalikit, C. (1984). The merits of multiple-answer items as evaluated by using six scoring formulas. *Journal of Experimental Education, 52*, 152-158.

Jansen, P.G.W., and Roskam, E.E. (1986). Latent trait models and dichotomization of graded responses. *Psychometrika, 51*, 69-91.

Kessler, D. P. (April, 1988). Machine-scored versus grader-scored quizzes - An experiment. *Engineering Education*, 705-709.

Kirst, M.W. (1991). Interview on assessment issues with Lorrie Shepard. Educational Researcher, 20, 21-23.

Kuhn, T.S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Levine, M., & Drasgow, F. (1983). The relation between incorrect option choice and estimated ability. *Educational and Psychological Measurement, 43*, 675-685.

Linacre, J. M. (1989). FACETS: Many-faceted Rasch analysis with FACFORM data formatter. (Program manual.) Chicago: MESA Press.

MacRae, D.J., & Green, D.R. (1971). Comment on Fry's "Orangoutang Score." *The Reading Teacher, 24*, 364.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.

Masters, G.N. (1988a). Measurement models for ordered response categories. Chapter 1 in **Latent trait and latent class models**, R. Langeheine & J. Rost (Eds.). New York: Plenum Press.

Masters, G. N. (1988b). The analysis of partial credit scoring. *Applied Measurement in Education, 1*, 279-298.

Masters, G.N., & Wright, B.D. (1984). The essential process in a family of measurement models. *Psychometrika, 49*, 529-544.

O'Connor, J. (1940). Unsolved business problems. Boston: Johnson O'Connor Research Foundation.

Resnick, L. B., & Resnick, D. P. (in press). Assessing the thinking curriculum: New tools for educational reform. To appear in Gifford, B. R., and O'Connor, M. C. (Eds.), **Future assessments: Changing views of aptitude, achievement, and instruction**. Boston: Kluwer Academic Publishers.

Rosenshine, B. (1980). Skill hierarchies in reading comprehension. In Shapiro, R.J., Bruce, B. C., & Brewer, W. F. (Eds.), **Theoretical issues in reading comprehension**. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sax, G. (1962). Theoretically derived chance scores and their normative equivalents on a selected number of standardized tests. Educational and Psychological Measurement, 22, 573-576.

Smith, R. M. (1987). Assessing partial knowledge in vocabulary. *Journal of Educational Measurement, 24*, 217-231.

Stiggins, R. J., & Bridgeman, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*, 271-286.

Thissen, D. J. (1976). Information in wrong responses to Raven's Progressive Matrices. *Journal of Educational Measurement, 13*, 201-214.

Wright, B. D., & Linacre, J. M. (1989). The differences between scores and measures. *Rasch Measurement SIG Newsletter, 3*, 1-4.

Wright, B. D., & Masters, G. N. (1982). **Rating Scale Analysis**. Chicago: MESA Press.

Yan, J. (1989). Differential option weighting for multiple-choice test items. Unpublished paper.

Ziller, R.C. (1957). A measure of the gambling response-set in objective tests. *Psychometrika, 22*, 289-292.

Figure 1(a). Item Characteristic Curves

6 + 8 = ? (Masters, 1988)

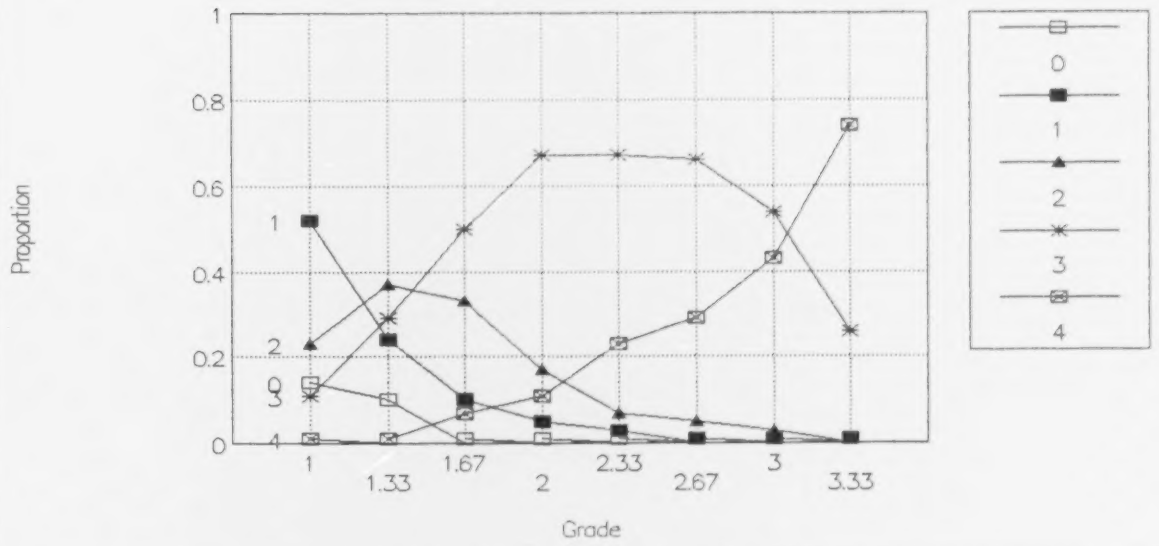


Figure 1(b). Item Response Map

6 + 8 = ? (Masters, 1988)

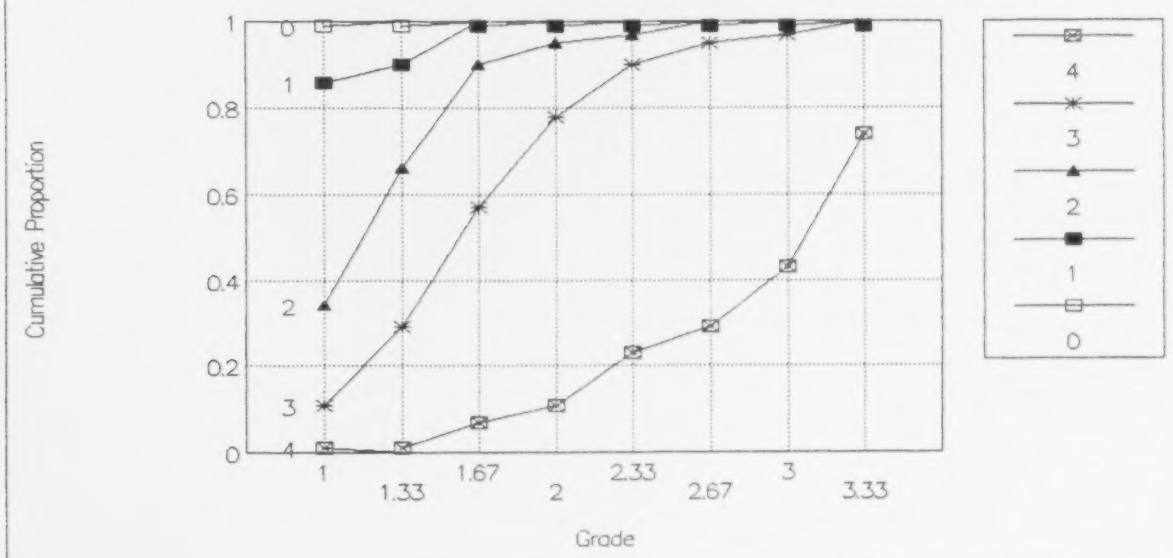


Figure 2(a): Masters' PCM

Item Characteristic Curves

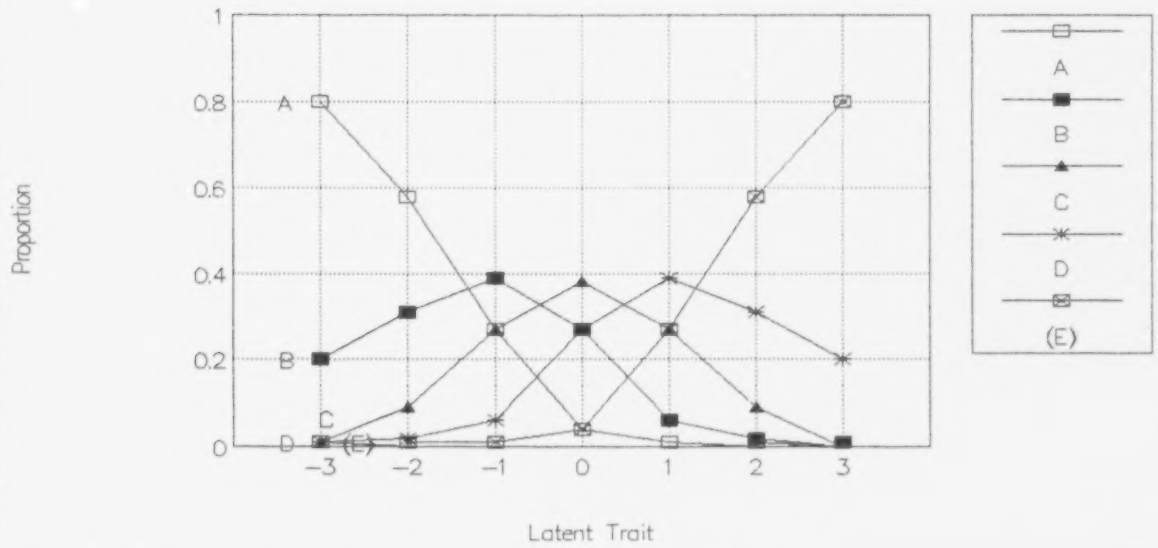


Figure 2(b): Masters' PCM

Item Response Map

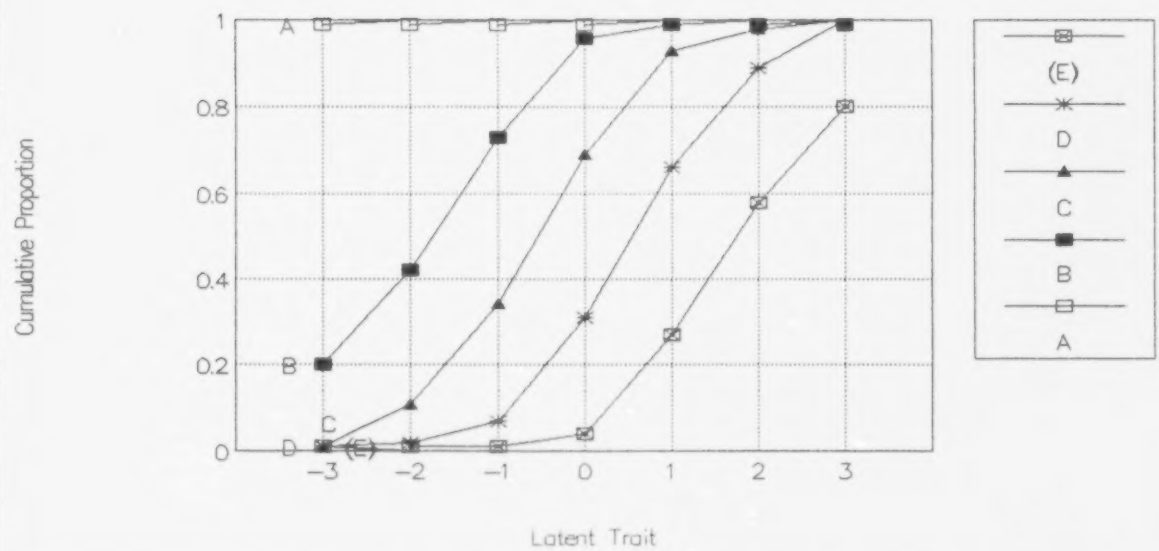


Figure 3(a): Jansen & Roskam Item Characteristic Curves

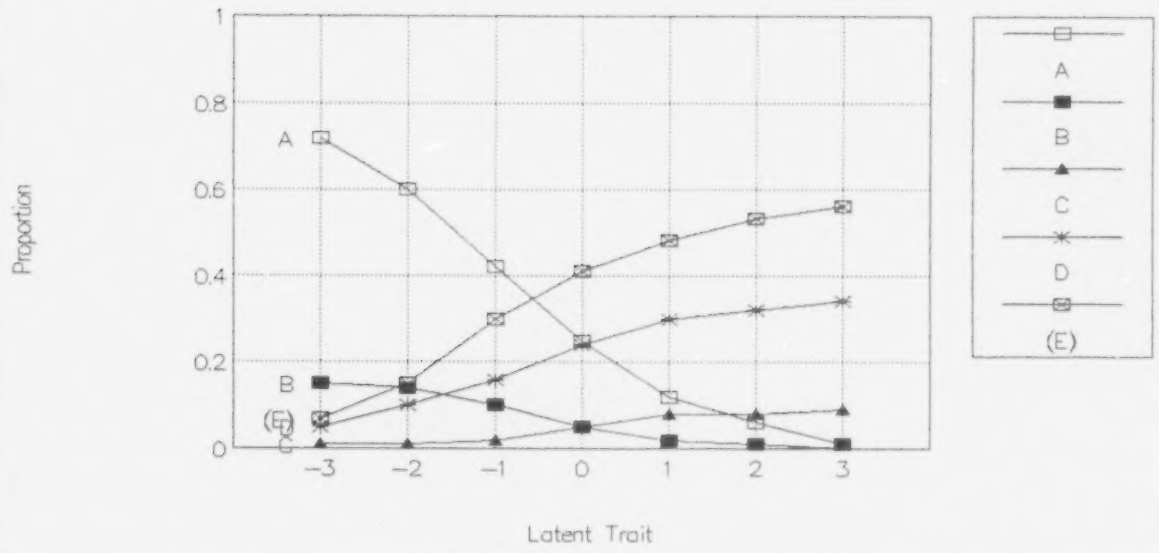


Figure 3(b): Jansen & Roskam Item Response Map

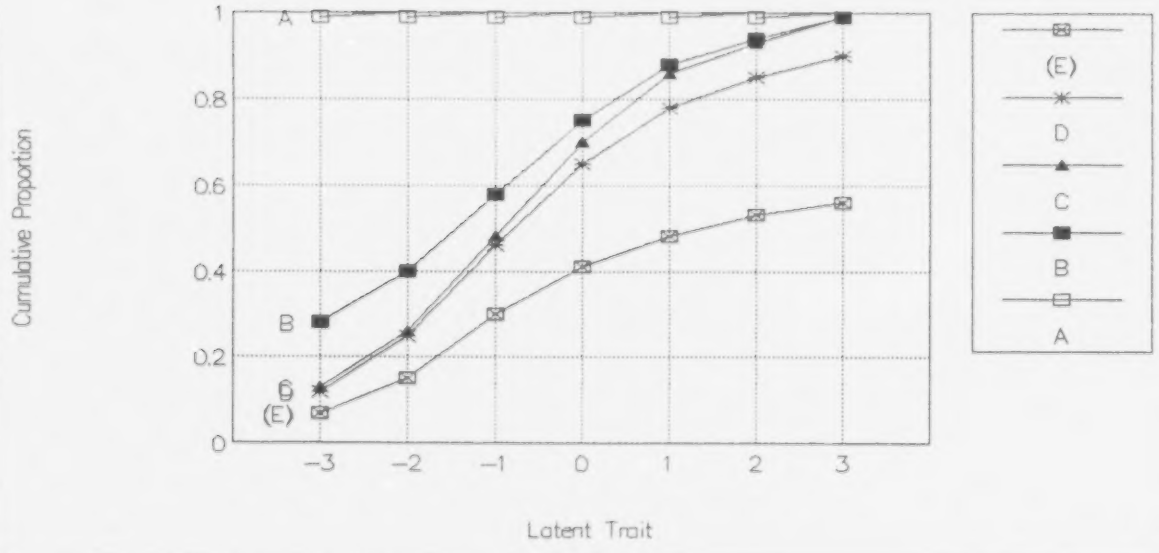


Figure 4(a). Item 17 (Section 1)

Estimate D = rt

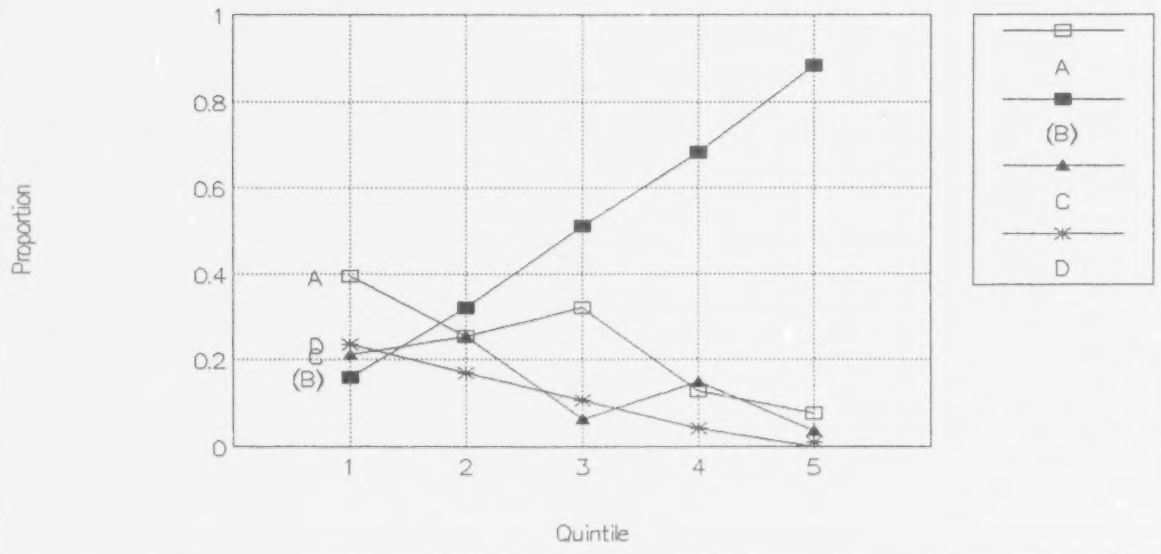


Figure 4(b). Item 17 (Section 1)

Estimate D = rt

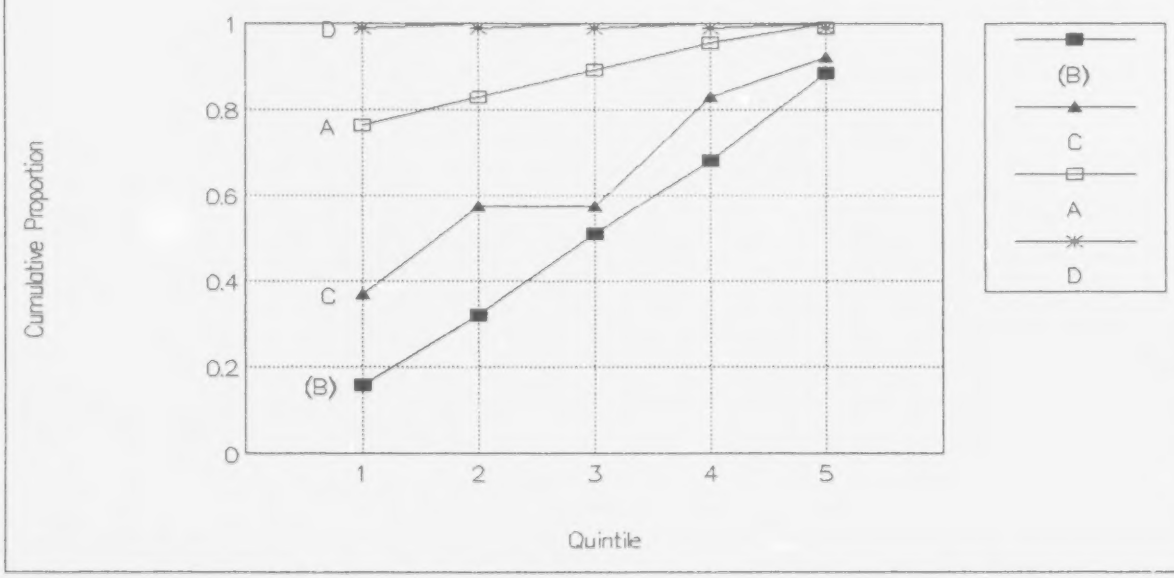


Figure 5(a). Item 18 (Section 1)

Square Root of 2501

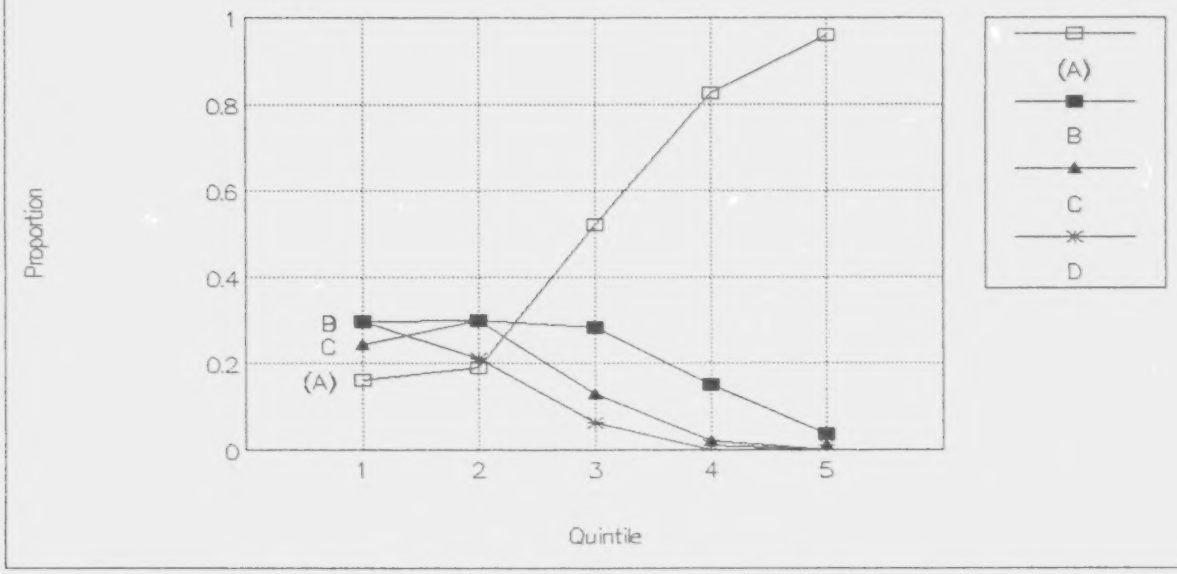


Figure 5(b). Item 18 (Section 1)

Square Root of 2501

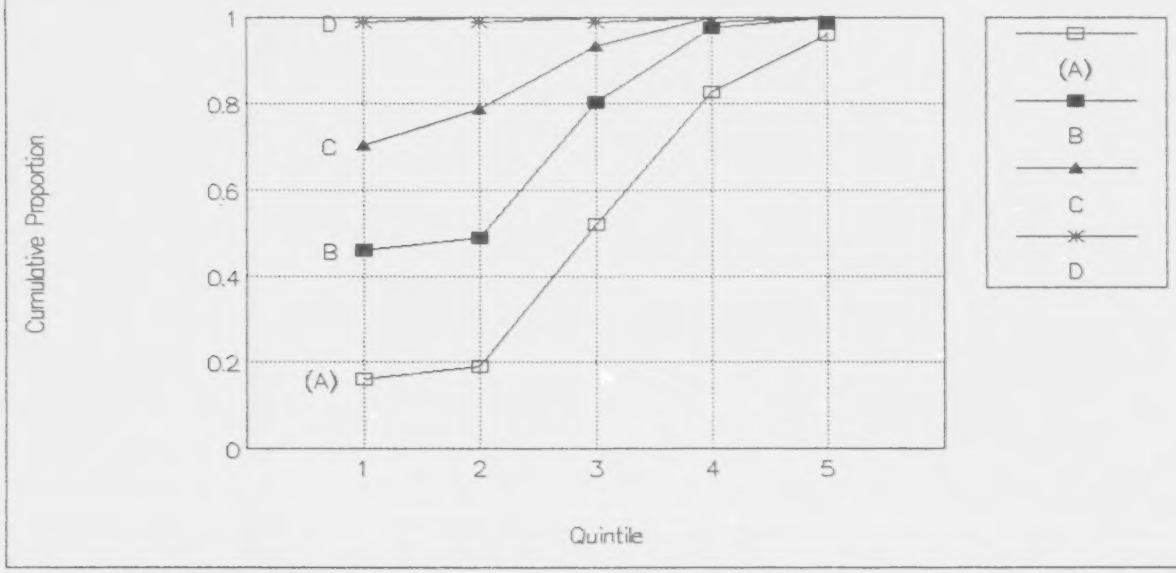


Figure 6(a): Item 10 (Sect 2)

Income/Expense Trend

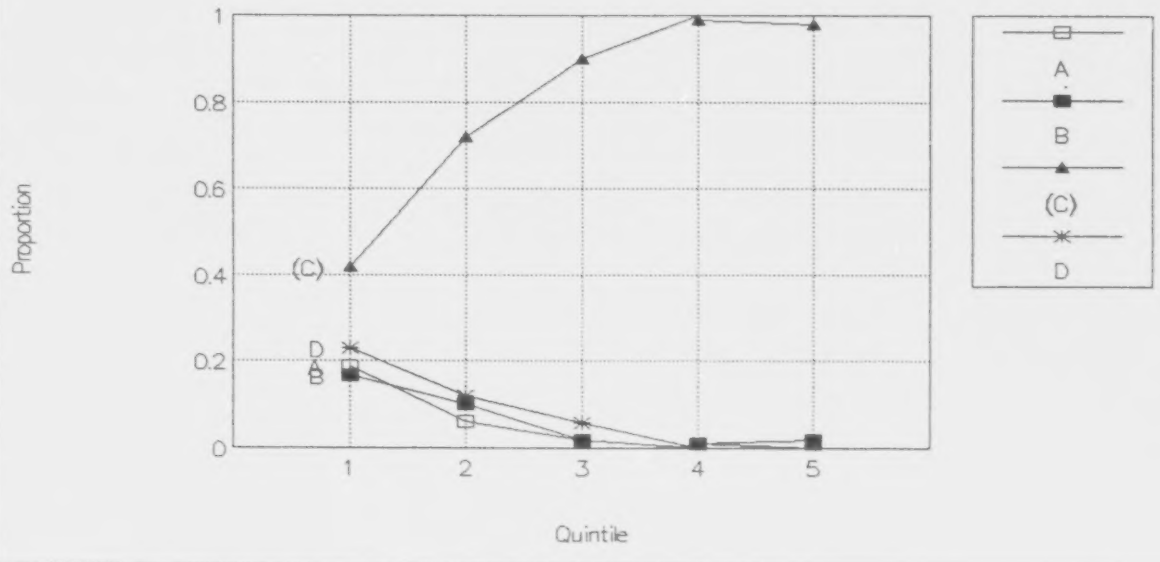


Figure 6(b): Item 10 (Sect 2)

Income/Expense Trend

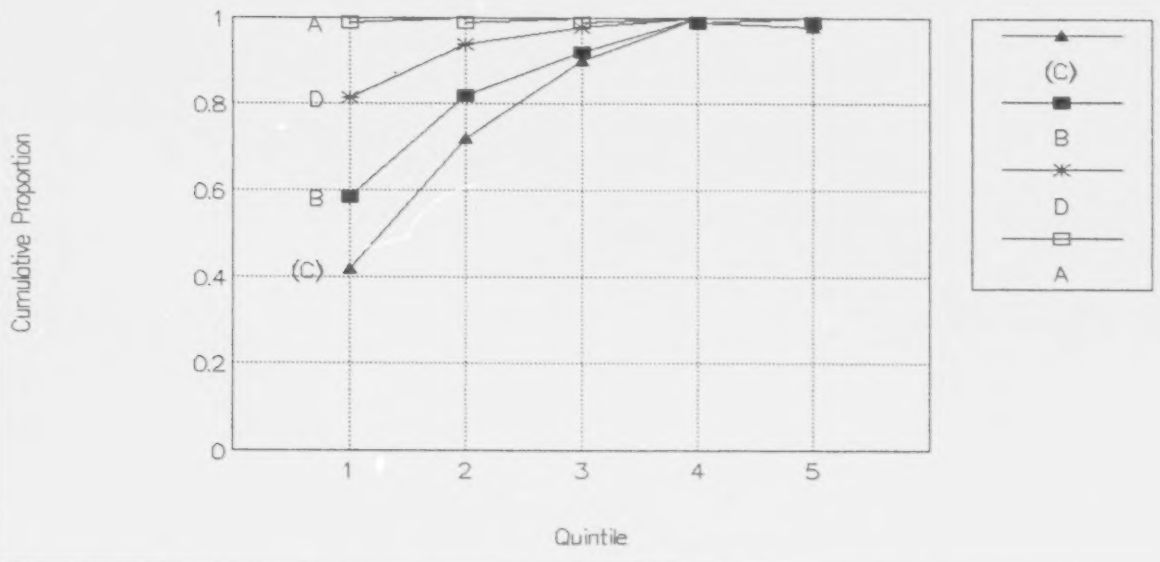


Figure 7(a): Item 15 (Sect 2)

Sum of (−3, −15)

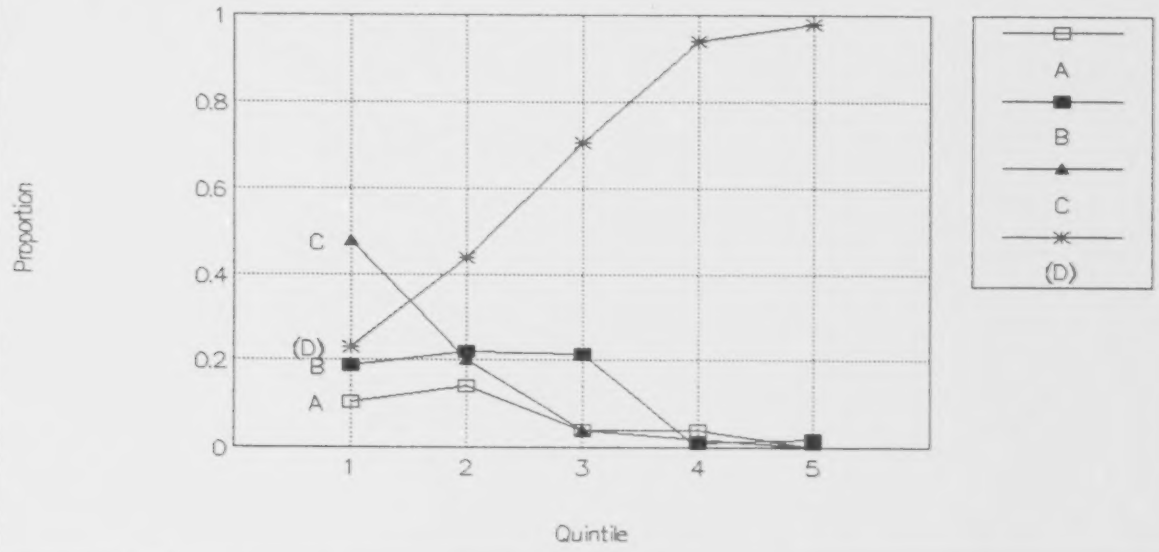


Figure 7(b): Item 15 (Sect 2)

Sum of (−3, −15)

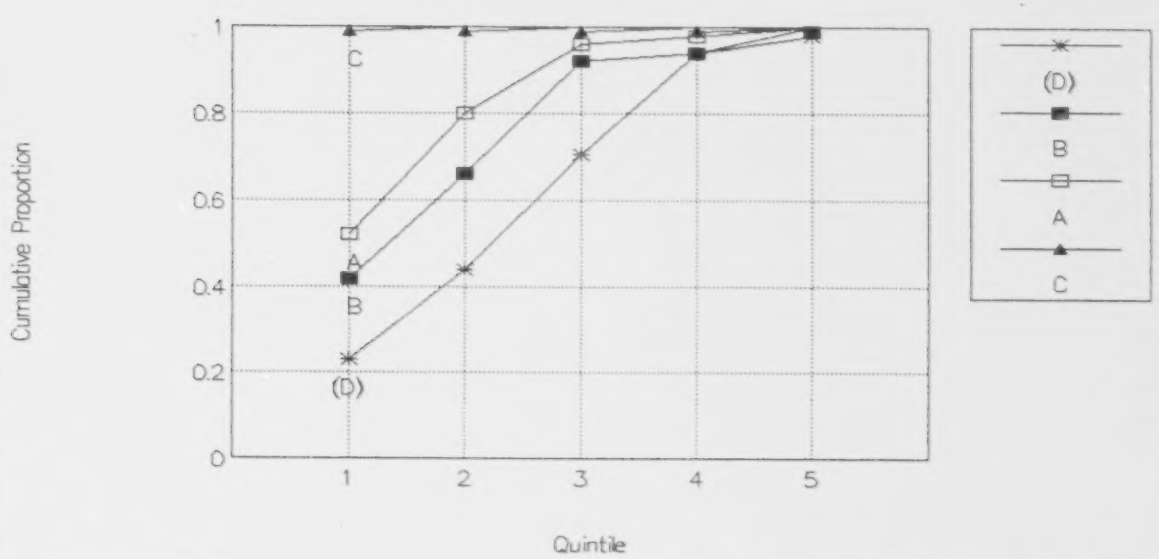


Figure 8(a): Item 16 (Sect 2)

Ordered Pairs

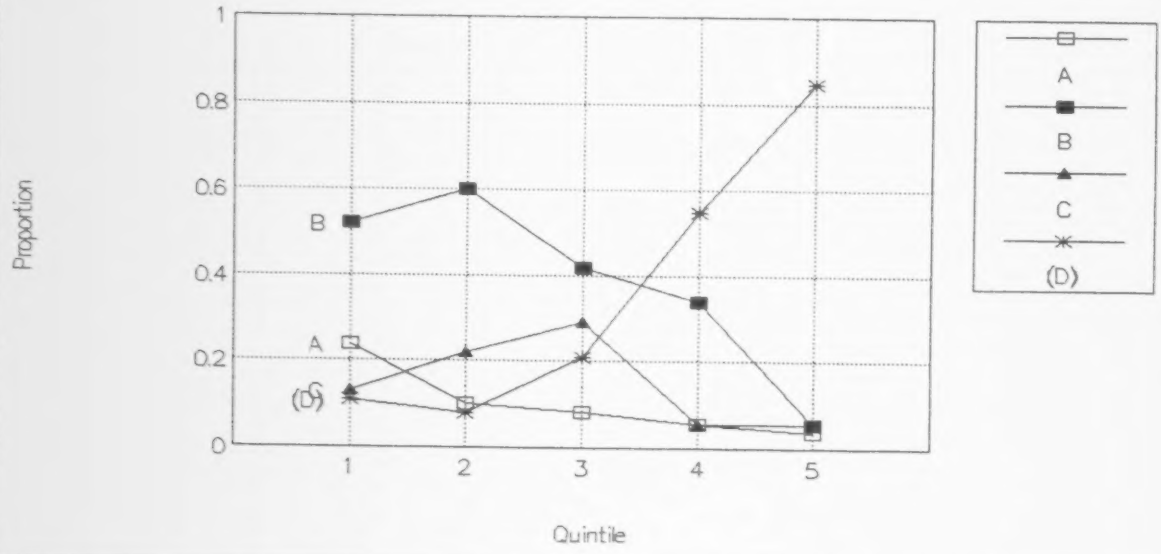


Figure 8(b): Item 16 (Sect 2)

Ordered Pairs

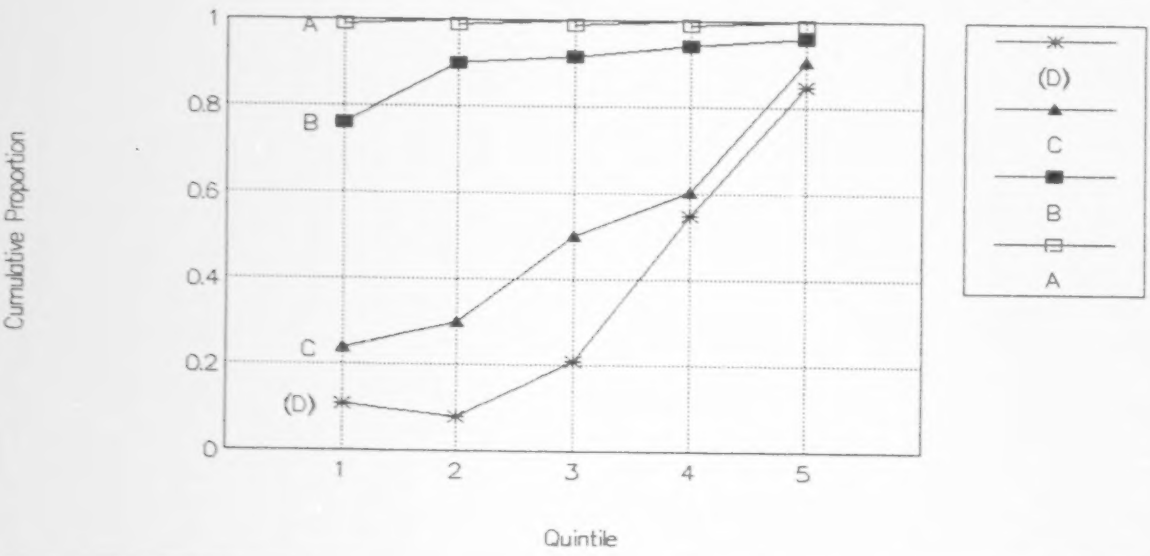


Figure 9(a). Item 2 (Section 3)

Which number has GREATEST value?

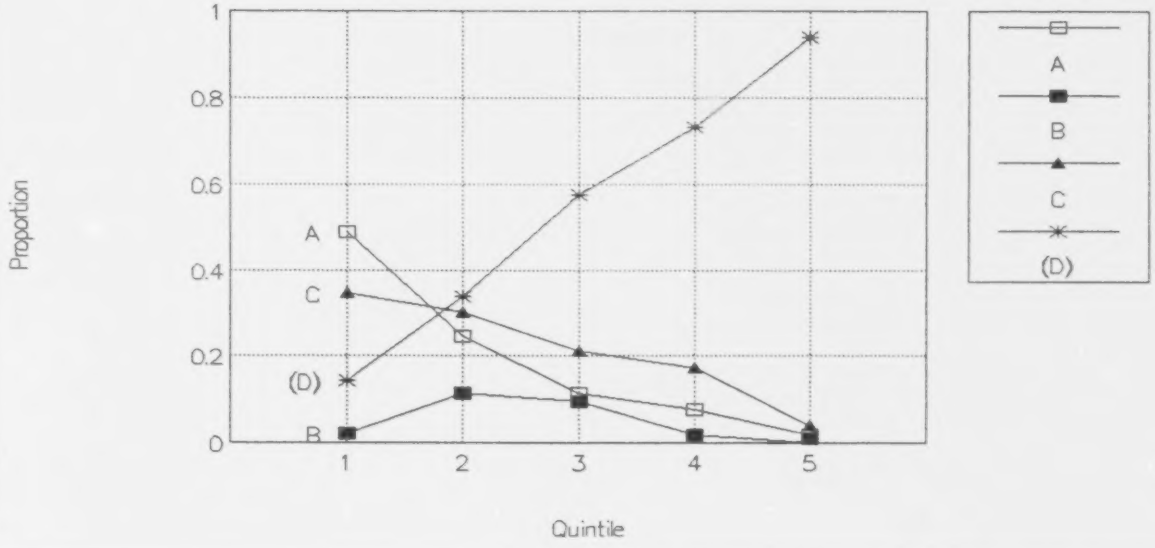


Figure 9(b). Item 2 (Section 3)

Which number has GREATEST value?

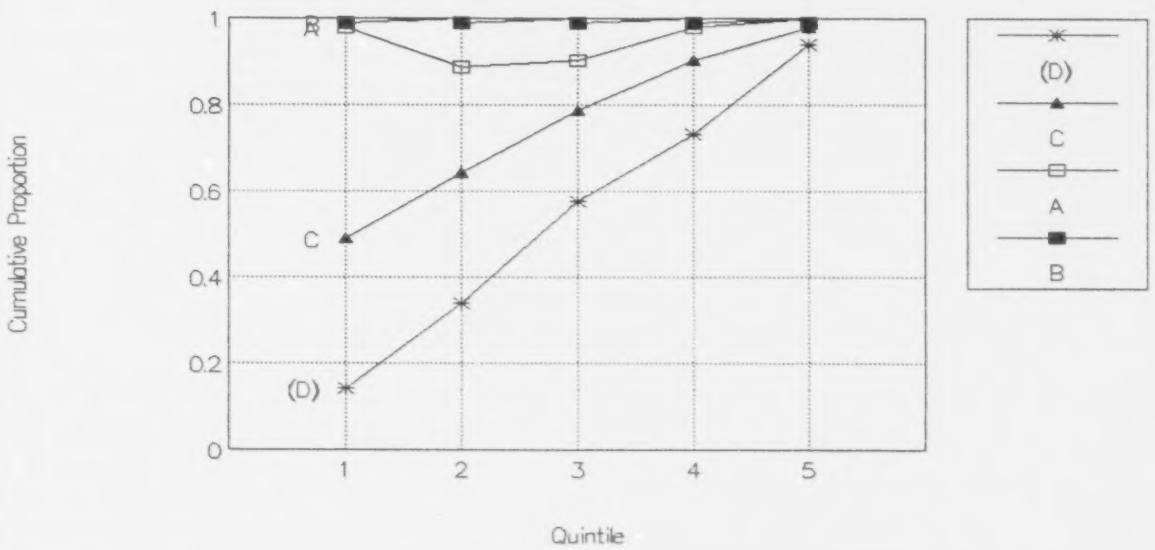


Figure 10(a): Item 35 (Sect 3)

Filling Tank w/Pail

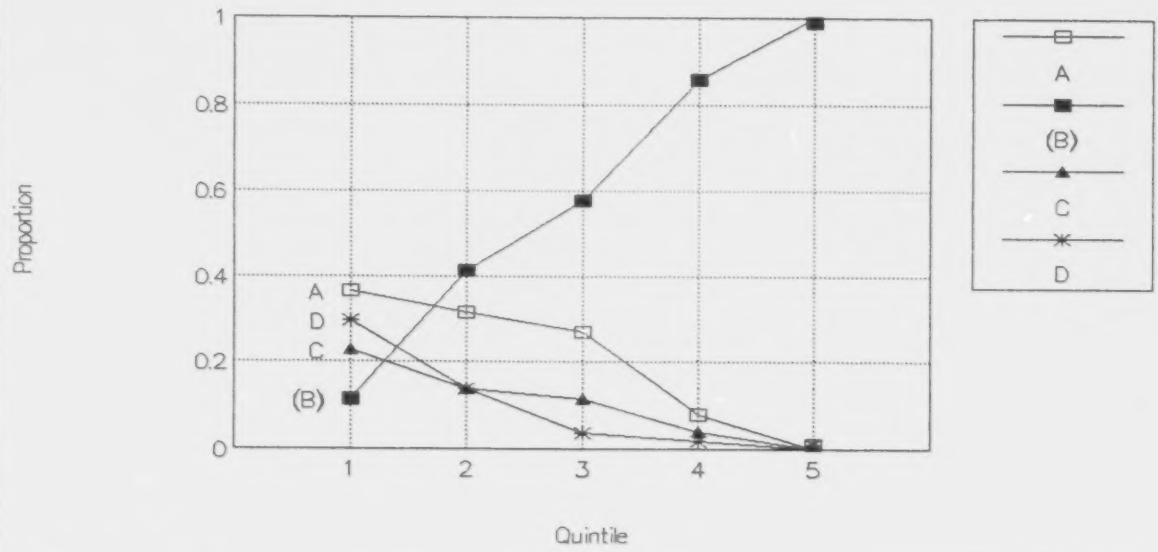


Figure 10(b): Item 35 (Sect 3)

Filling Tank w/Pail

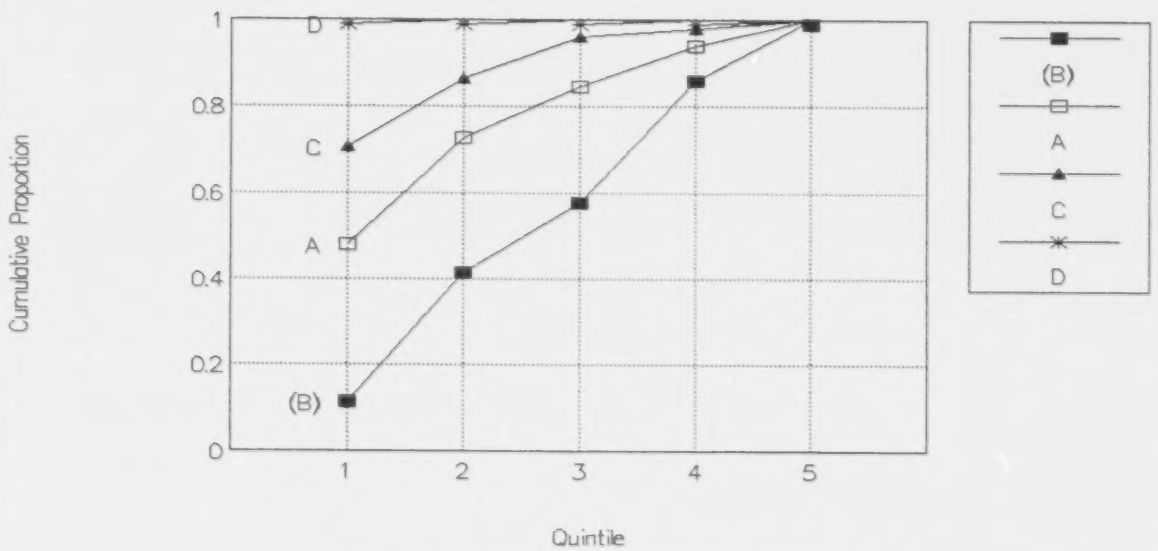


Figure 11(a). Item 45 (Section 3)

Hexagonal Pyramid

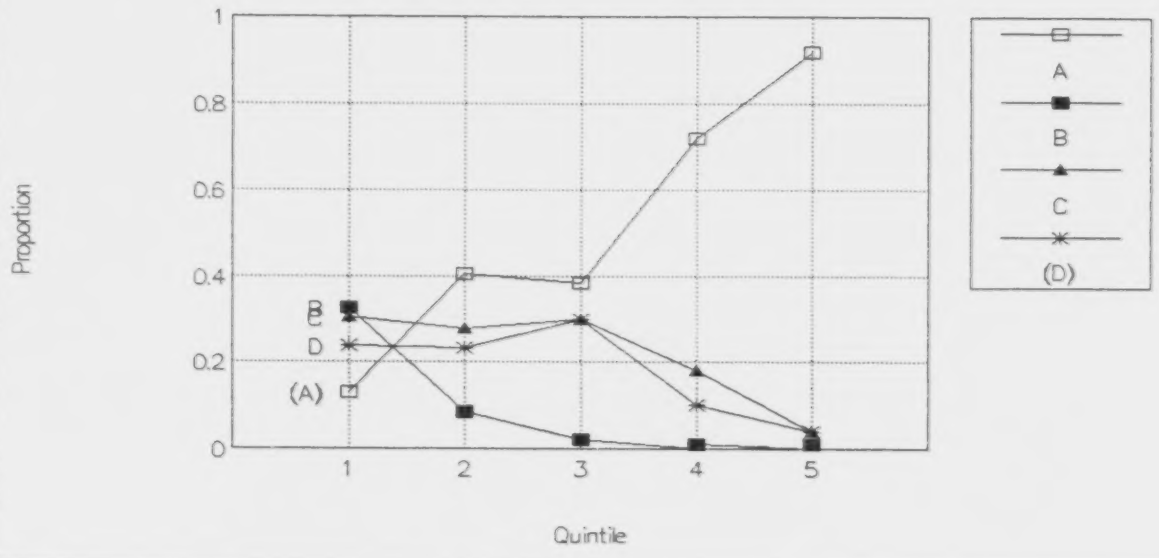


Figure 11(b). Item 45 (Section 3)

Hexagonal Pyramid

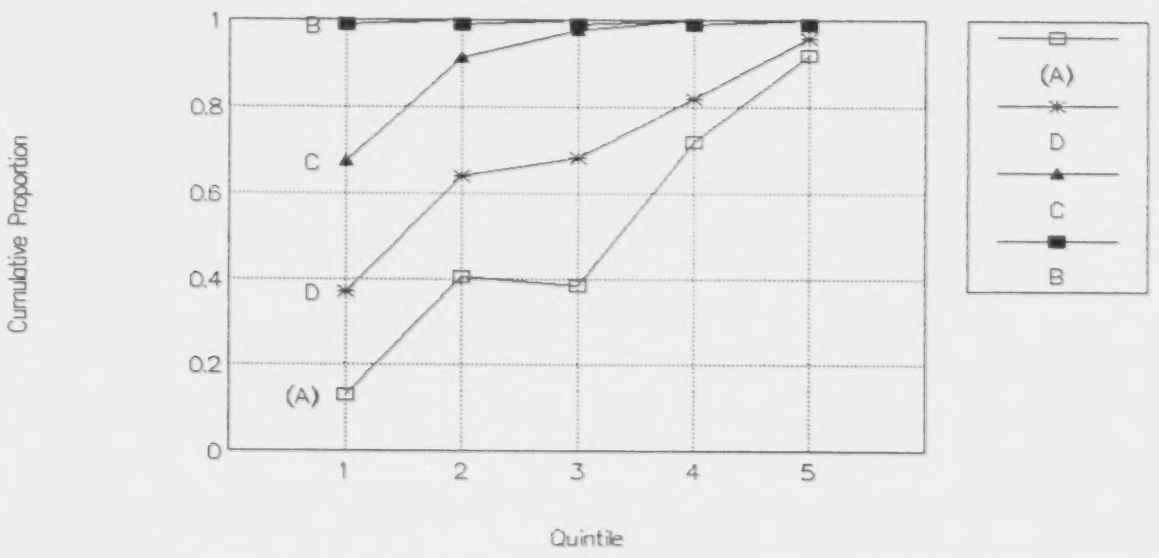


Figure 12(a). Item 56 (Section 3)

Heads on three coin tosses

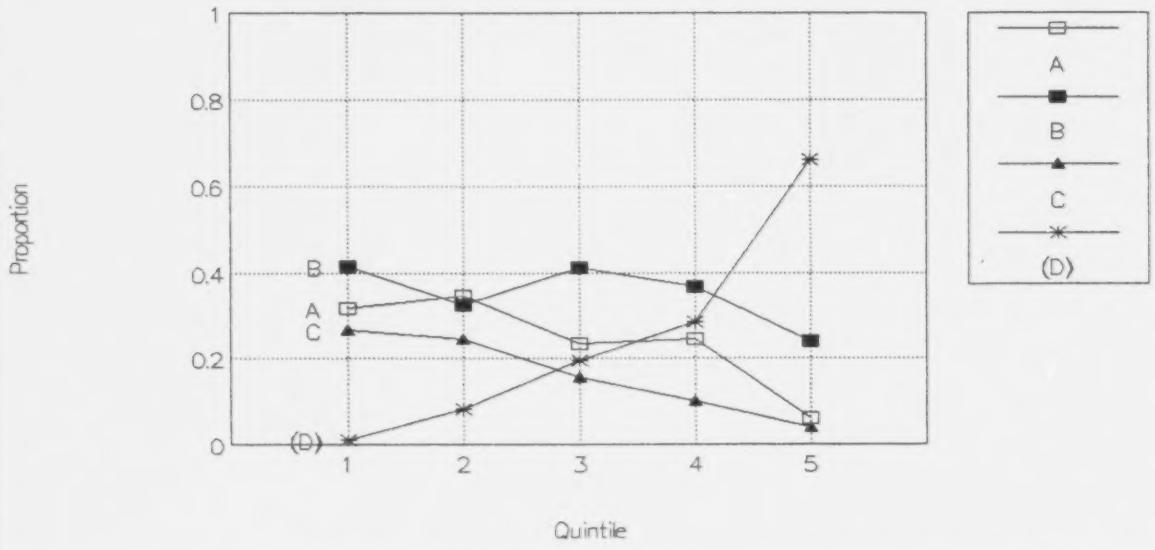


Figure 12(b). Item 56 (Section 3)

Heads on three coin tosses

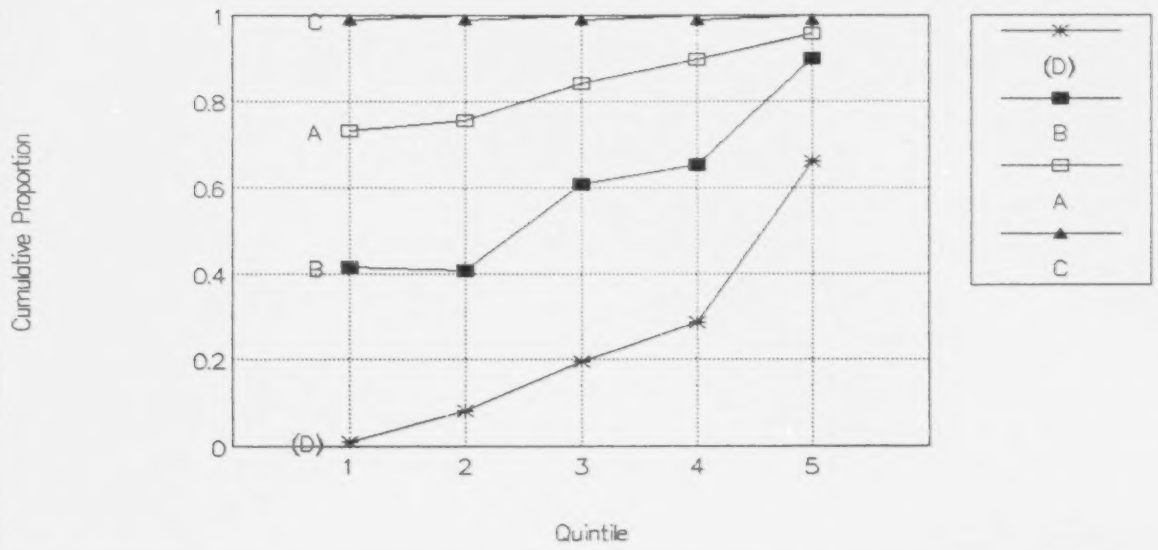


Figure 13(a): Item 62 (Sect 3)

Area Problem

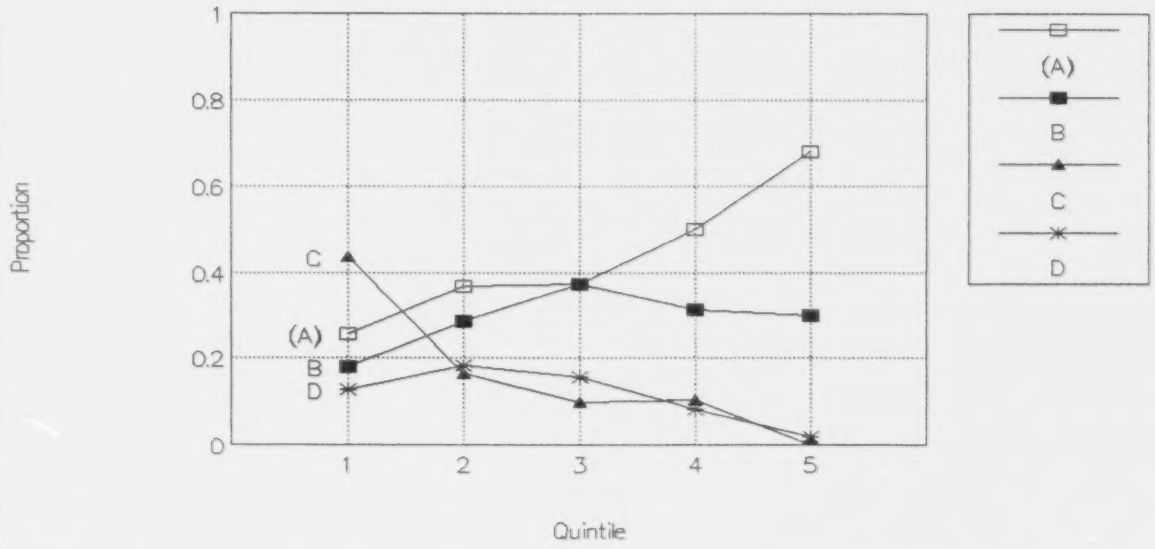


Figure 13(b): Item 62 (Sect 3)

Area Problem

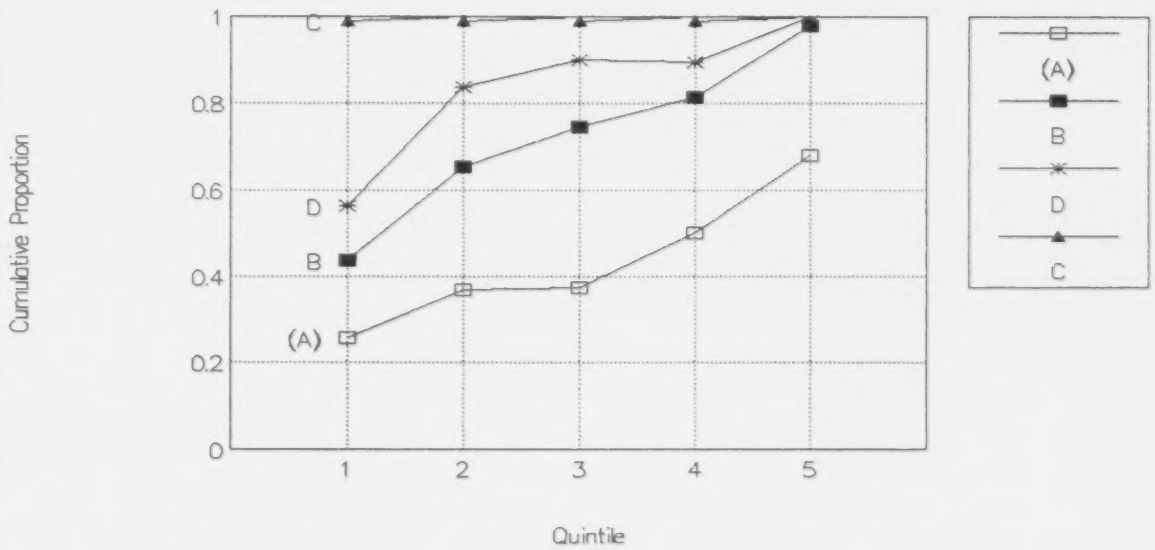


Figure 14(a): Item 69 (Sect 3)

Iron and Water in Soil

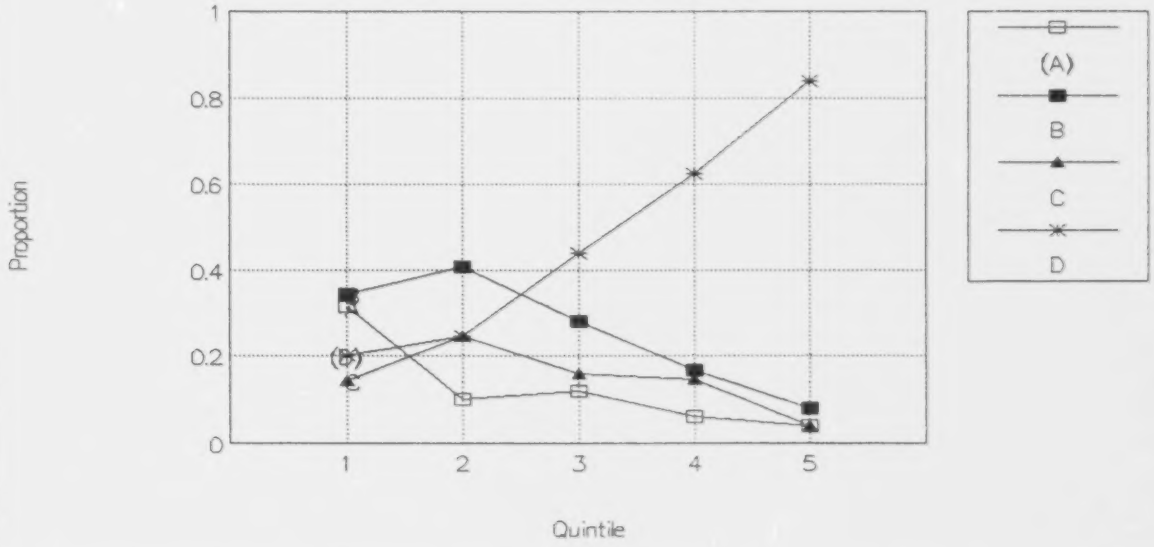


Figure 14(b): Item 69 (Sect 3)

Iron and Water in Soil

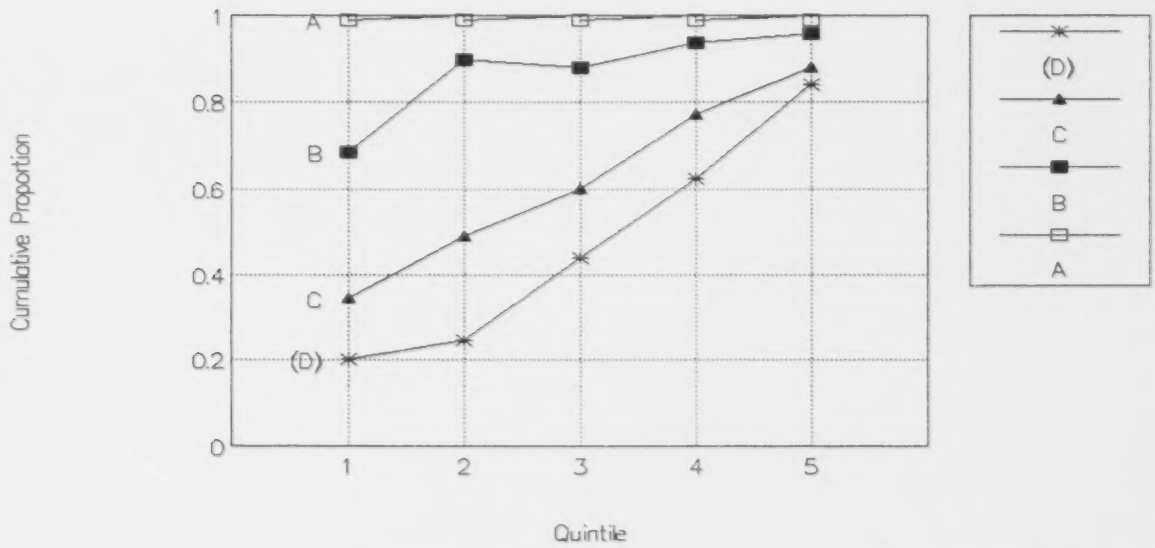


Figure 15(a). Item 91 (Section 3)

Hydrogen Peroxide

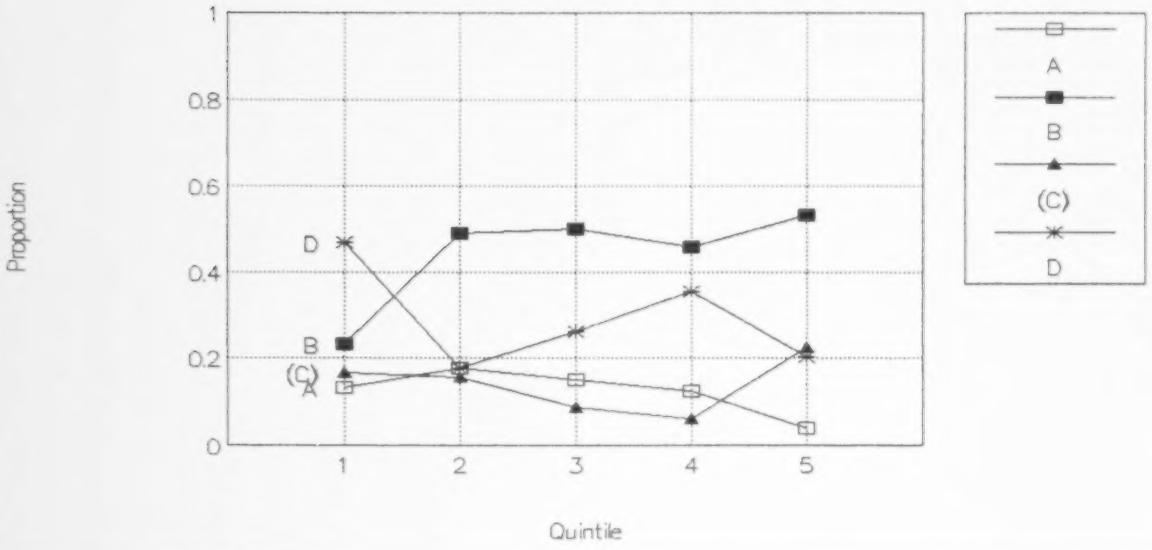


Figure 15(b). Item 91 (Section 3)

Hydrogen Peroxide

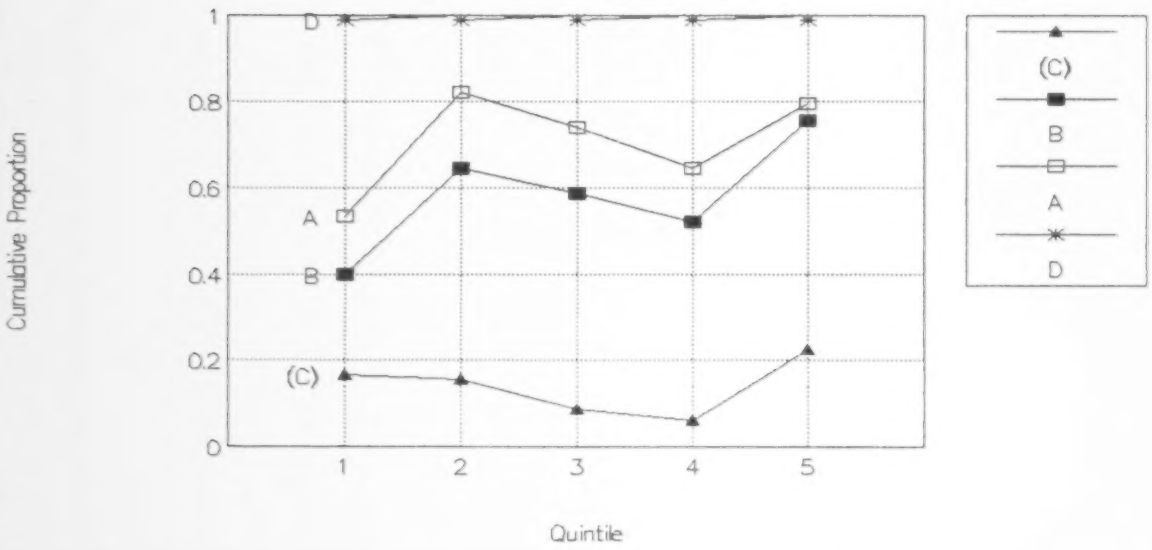


Figure 16(a). Item 92 (Section 3)

Robot Selection

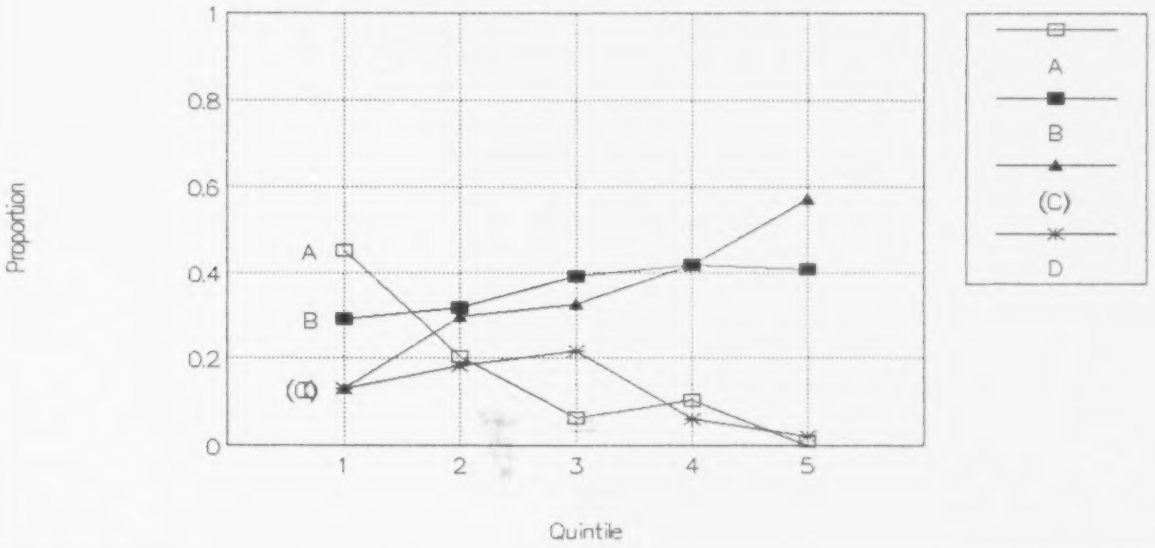


Figure 16(b). Item 92 (Section 3)

Robot Selection

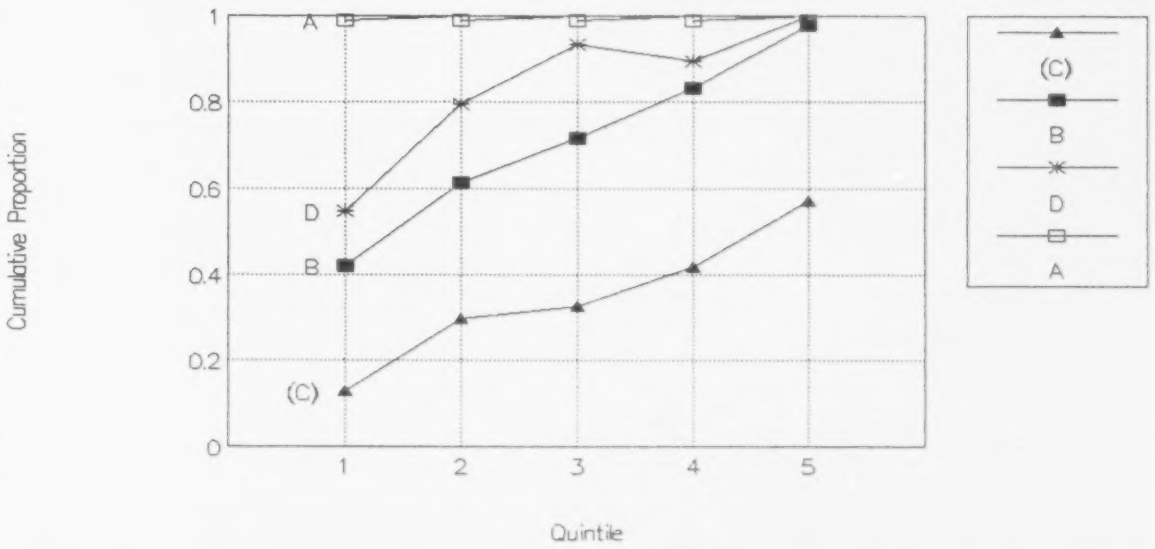


Figure 17(a): Item 93 (Sect 3)

Public Opinion Poll

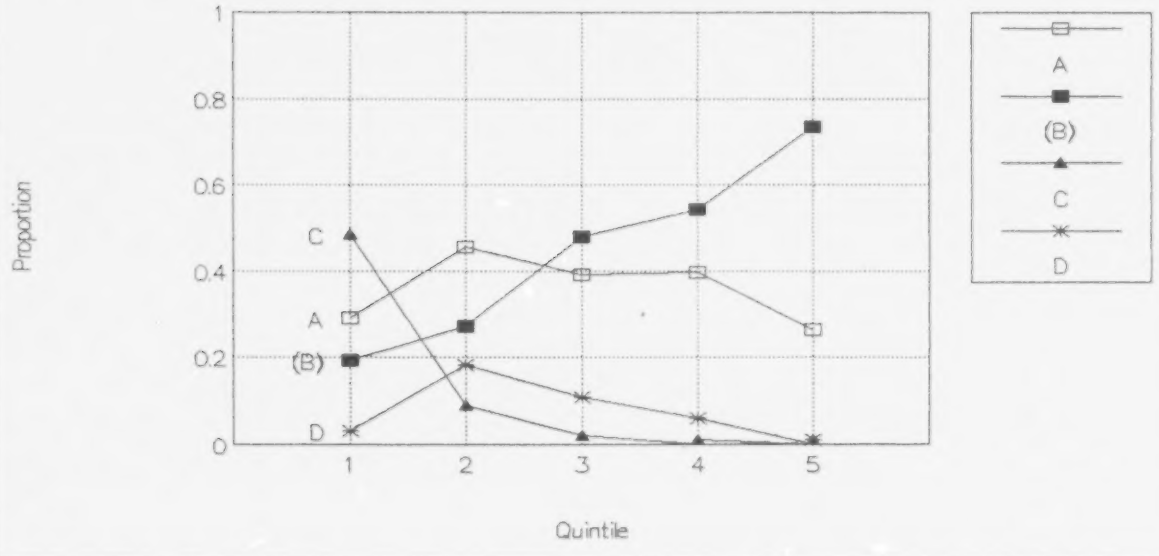


Figure 17(b): Item 93 (Sect 3)

Public Opinion Poll

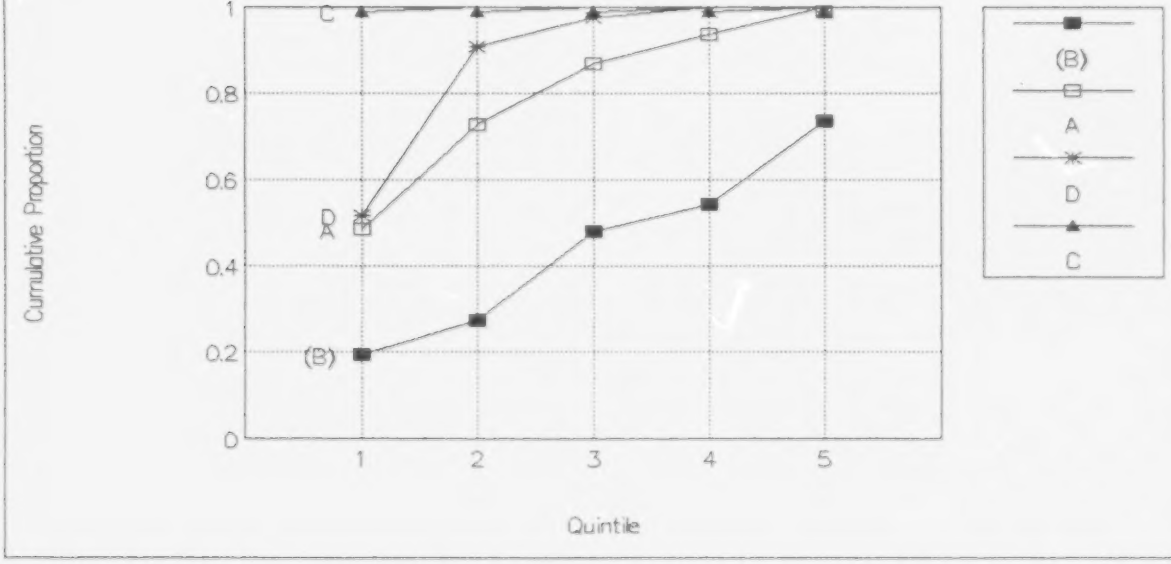


Figure 18(a). Item 94 (Section 3)

Meal Preparation

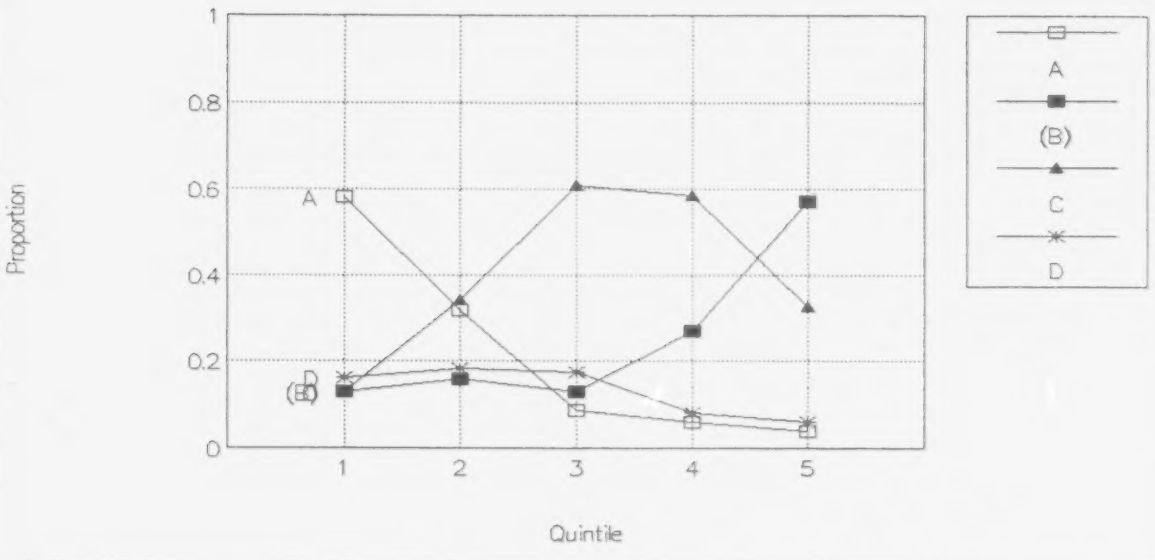


Figure 18(b). Item 94 (Section 3)

Meal Preparation

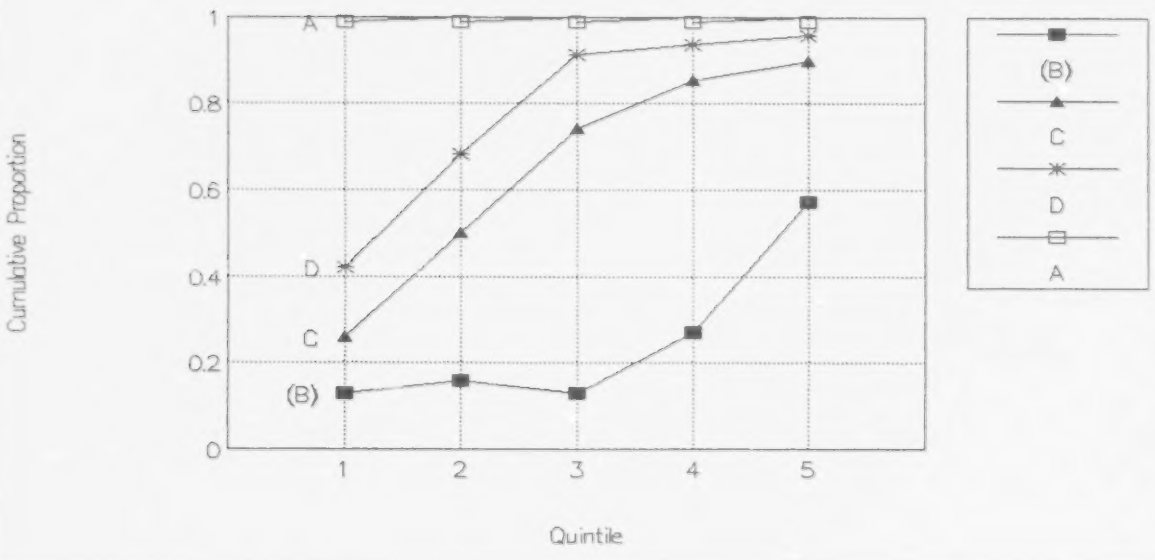


Figure 19(a): Item 96 (Sect 3)

Airlock Wear

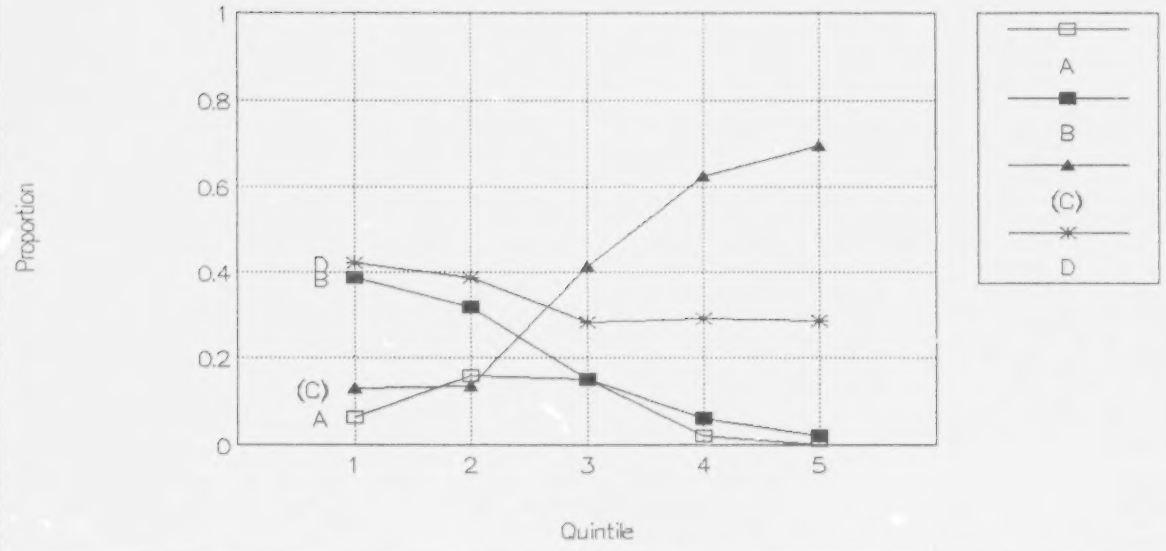
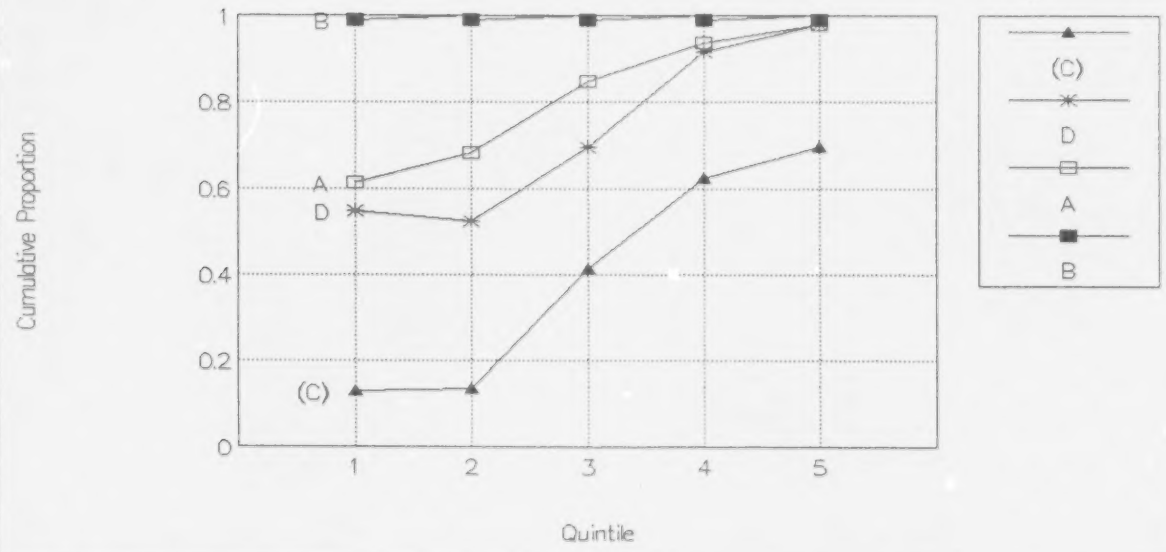


Figure 19(b): Item 96 (Sect 3)

Airlock Wear



DIRECTIONS: Find the BEST answer.

17. If r = 75 and t = 4.95, find the BEST ESTIMATE of D = rt.

A 400

B 375

C 300

D 280

18. $\sqrt{2501}$ is APPROXIMATELY

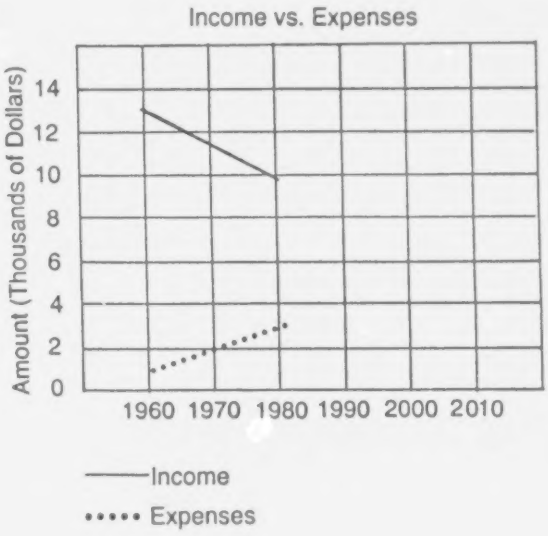
A 50

B 500

C 1250

D 2500

10. If the present trends continue, what will be the expenses in 2010?



A $3,000

B $4,000

C $6,000

D $9,000

15. −3 − 15 =

A 18

B 12

C −12

D −18

DIRECTIONS: Find the BEST answer.

16. Which of these sets of ordered pairs is described by $y - x^3$?

A {(1,2), (2,4), (3,6)}

B {(1,3), (2,6), (3,9)}

C {(1,1), (2,4), (3,9)}

D {(1,1), (2,8), (3,27)}

2. Which of these numbers has the GREATEST value?

A 6.575×10^{-3}

B 1.27×10^{-2}

C 5.40×10^{-4}

D 5.01×10^{-1}

35. Tony was filling a tank with a 2.4 L pail. After putting in 6 full pails of water, the tank was one-fourth full. How much water will be in the tank when it is full?

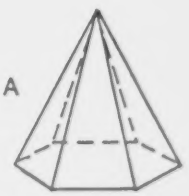
A 24 L

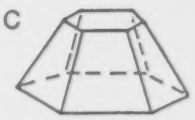
B 57.6 L

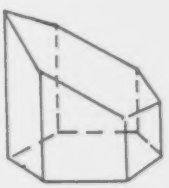
C 72 L

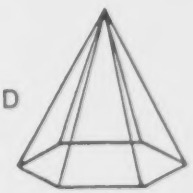
D 86.4 L

45. Which of these is an appropriate drawing of a hexagonal pyramid?

A 

C 

B 

D 

56. What is the probability of getting three heads in a row tossing a fair coin?

A $\frac{3}{2}$

B $\frac{1}{2}$

C $\frac{1}{4}$

D $\frac{1}{8}$

Background Information

Questions 91 through 96 involve the inhabitants of a distant planet. This planet, called Romula, is located in a different galaxy. The Romulans have built the Alpha Centauri Space Station (ACSS), which is a self-sufficient space capsule. It is staffed and populated only by people from Romula and their families. Many have compared the ACSS with a floating city. Most services are provided within the space station. Try to imagine what it would be like to live and work on such a space station as you answer these questions.

91. The ACSS has a factory which decomposes hydrogen peroxide to produce water and oxygen for the station's living areas. Only oxygen and water are produced. There are no by-products or waste from this process. The ratio (by mass) of oxygen to water produced is 16 to 18. An emergency supply of 160,000 kg of oxygen is needed immediately. If each supply vehicle can carry 4,000 kg of hydrogen peroxide, how many vehicles must be brought into the decomposition plant to supply enough hydrogen peroxide to produce the needed oxygen?

A 9 vehicles

B 40 vehicles

C 85 vehicles

D need more information

92. A Romulan family needs a robot to do cleaning tasks three times per week, repair tasks twice per week, and aluminum recycling tasks once per week. It is important that all work be completed in a minimum amount of time. Which of the three available robots from the chart below should be selected?

A Robot 347

B Robot 48A

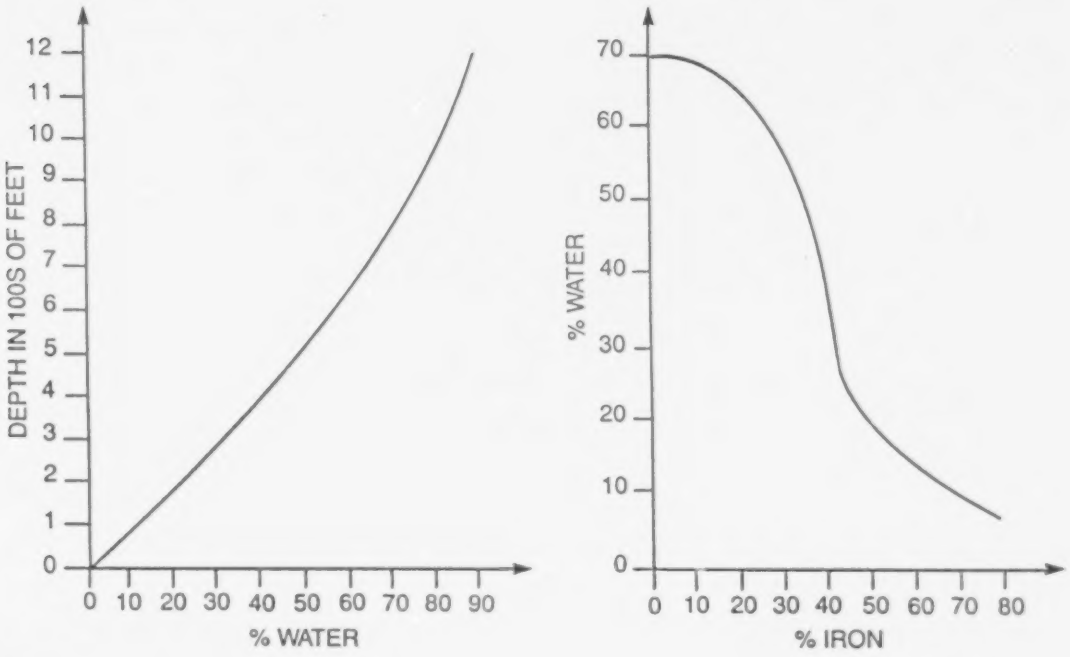
C Robot 50C

D cannot be determined

	Average Completion Time per Task		
Robot	**Repairs**	**Cleaning**	**Recycling**
Robot 347	1 hour	2 hours	1 hour
Robot 48A	45 minutes	1.5 hours	30 minutes
Robot 50C	0.5 hours	1 hour	1.5 hours

DIRECTIONS: Find the BEST answer.

Data on iron and water composition from soil sample analyses were graphed:



69. Based on the information in the graphs, which of the following statements is true?

A The greater the depth, the less percentage of water in the sample.

B The greater the percentage of water in the sample, the greater the percentage of iron in the sample.

C The greater the depth, the greater the percentage of iron in the sample.

D The greater the depth, the less percentage of iron in the sample.

DIRECTIONS: Find the BEST answer.

93. A public opinion firm has been asked to determine what the Romulans think about the ACSS. To get the most representative sample, whom should they interview?

A Romulans living on the ACSS

B Romulans living on the ACSS and/or on the planet

C political leaders from the planet

D Romulans living on the planet

94. On the ACSS, computers are programmed to prepare the meals. The evening meal is to be served at 5:30 p.m. It consists of a casserole which takes 30 minutes to cook and must cool 10 minutes before serving; bread which is to be warmed for 35 minutes; and vegetables which must cook for 15 minutes. There is also a dessert to be served at 6:00 p.m which must cook for 60 minutes. In which order should the computer prepare the food?

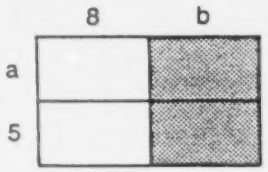
A casserole, bread, vegetables, dessert

B casserole, bread, dessert, vegetables

C dessert, casserole, bread, vegetables

D dessert, bread, casserole, vegetables

62. If the area of a rectangle is found by multiplying length by width, find the area of the shaded region below.



A ab + 5b

B 5ab

C ab + 5

D a + 5b

DIRECTIONS: Find the BEST answer.

96. On the ACSS, there is an airlock which permits the people on the station to go in and out without losing air from the living/working quarters. Since the parts wear a little every time the airlock is opened and closed, it is necessary to keep track of the wear and replace the parts when they become too worn. If the wear reaches Level 5, serious problems will occur, bringing danger to the ACSS crew and their families. What is the latest stardate when corrective action should be taken, based on the chart below?

FIGURE 1. Airlock Wear Level as a Function of Time

A 2

B 3

C 7

D 8

MEAP MATH; SHORT VERSION GRADE 10 "BIGSCALE" RASCH ANALYSIS VER. 1.73 Jan 31 20:25:51 1991
INPUT: 262 PERSONS 16 ITEMS ANALYZED: 249 PERSONS 16 ITEMS 2 CATEGORIES TABLE B-1

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11 3476218140653952	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
85		0	0	OUT					
256	1111111010000000	8	16	-.01	.55	.6	-2.2	.6	-1.6
171	1111111000100000	8	16	-.01	.55	.7	-1.9	.6	-1.5
224	1111111010100000	9	16	.30	.56	.7	-1.8	.6	-1.4
245	11 1111110100000	9	15	.55	.58	.6	-1.8	.6	-1.3
251	1111111110100000	10	16	.61	.57	.6	-1.8	.5	-1.4
162	1111101011000000	8	16	-.01	.55	.7	-1.7	.6	-1.4
242	1111111011100000	10	16	.61	.57	.7	-1.6	.6	-1.3
145	1110111111000000	9	16	.30	.56	.7	-1.5	.6	-1.2
154	1101111100000000	7	16	-.31	.55	.7	-1.5	.6	-1.2
172	1111011100100000	8	16	-.01	.55	.7	-1.5	.6	-1.3
183	1111110110010000	9	16	.30	.56	.7	-1.5	.6	-1.2
208	1111100000000000	5	16	-.96	.59	.6	-1.5	.5	-1.2
223	1111111111010000	11	16	.95	.59	.6	-1.5	.5	-1.2
77	10	1	2	-.74	1.42	.9	-1.4	.9	-.1
161	1111100100100000	7	16	-.31	.55	.7	-1.4	.7	-1.1
247	1111111111110000	12	16	1.32	.63	.6	-1.4	.4	-1.1
32	1111101111010000	10	16	.61	.57	.7	-1.3	.6	-1.0
59	1111100110100000	8	16	-.01	.55	.8	-1.3	.7	-1.0
229	11111001001 0000	7	15	-.21	.57	.8	-1.3	.7	-1.0
33	1111101011010000	9	16	.30	.56	.8	-1.2	.7	-1.1
86	11 00 0	2	5	-1.26	1.00	.5	-1.2	.5	-1.2
102	10	1	2	-1.77	1.46	.6	-1.2	.6	-.8
169	1111000000000000	4	16	-1.33	.63	.6	-1.2	.5	-1.0
219	1111000000000000	4	16	-1.33	.63	.6	-1.2	.5	-.9
97	11 000 0	2	6	-1.55	.93	.6	-1.1	.5	-1.0
150	1111011011010000	9	16	.30	.56	.8	-1.1	.7	-1.0
191	1111101110110000	10	16	.61	.57	.7	-1.1	.7	-.9
234	1111111111111000	13	16	1.76	.70	.6	-1.1	.4	-1.0
94	1110010100000000	5	16	-.96	.59	.7	-1.0	.6	-.8
144	1111111011011000	11	16	.95	.59	.7	-1.0	.6	-.8
163	1111100110001000	8	16	-.01	.55	.8	-1.0	.7	-.9
186	1111111101111000	12	16	1.32	.63	.7	-1.0	.5	-.8
209	1011111001000000	7	16	-.31	.55	.8	-1.0	.7	-.8
239	1111101100101000	9	16	.30	.56	.8	-1.0	.7	-.9
240	1110101000010000	6	16	-.62	.57	.8	-1.0	.7	-.8
10	1101000 0	3	8	-1.28	.78	.7	-.9	.7	-.7
78	11100111 0 0 0	6	11	-.19	.65	.8	-.9	.7	-.8
122	1111011010110000	9	16	.30	.56	.8	-.9	.7	-.9
127	1110101001100000	7	16	-.31	.55	.8	-.9	.7	-.9
140	1110110101010000	8	16	-.01	.55	.8	-.9	.7	-.9
182	1111111101000100	10	16	.61	.57	.8	-.9	.7	-.9
198	1110000100000000	4	16	-1.33	.63	.7	-.9	.6	-.7

MEAP MATH; SHORT VERSION GRADE 10 "BIGSCALE" RASCH ANALYSIS VER. 1.73 Jan 31 20:25:51 1991
INPUT: 262 PERSONS 16 ITEMS ANALYZED: 249 PERSONS 16 ITEMS 2 CATEGORIES TABLE B-1

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11								
	3476218140653952	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
233	1111010010001000	7	16	-.31	.55	.8	-.9	.8	-.7
237	1111010011010000	8	16	-.01	.55	.8	-.9	.7	-.9
80	1110 00 1 0 0	4	9	-.66	.73	.8	-.8	.7	-.8
180	1111101111111000	12	16	1.32	.63	.7	-.8	.7	-.5
236	1110010011000000	6	16	-.62	.57	.8	-.8	.7	-.7
249	1110111110101000	10	16	.61	.57	.8	-.8	.7	-.7
13	10 0 0	1	4	-2.26	1.25	.5	-.7	.4	-.4
43	1111111111110100	13	16	1.76	.70	.7	-.7	.5	-.7
146	1111 1100100100	8	14	.42	.60	.8	-.7	.7	-.7
152	1111 11110 0010	9	13	1.05	.67	.8	-.7	.6	-.7
188	1111101110111000	11	16	.95	.59	.8	-.7	.7	-.6
260	1110101 0	5	8	-.13	.77	.8	-.7	.7	-.6
37	1111111110110100	12	16	1.32	.63	.8	-.6	.6	-.6
44	1111111101110100	12	16	1.32	.63	.8	-.6	.6	-.6
50	1111111111111100	14	16	2.32	.81	.6	-.6	.3	-.6
55	1110000000001000	4	16	-1.33	.63	.8	-.6	.8	-.2
69	1101 0000010000	4	14	-1.06	.65	.8	-.6	.7	-.4
90	11001101 0 0 0	5	11	-.61	.65	.9	-.6	.8	-.6
143	1110110000101000	7	16	-.31	.55	.9	-.6	.8	-.6
167	1111111110110010	12	16	1.32	.63	.8	-.6	.6	-.6
178	1111111101110100	12	16	1.32	.63	.8	-.6	.6	-.6
184	1111111101010100	11	16	.95	.59	.8	-.6	.7	-.7
11	1011000 0	3	8	-1.28	.78	.8	-.5	.8	-.6
30	1111111110111100	13	16	1.76	.70	.8	-.5	.6	-.4
39	1111111110111100	13	16	1.76	.70	.8	-.5	.6	-.4
51	1111111111011010	13	16	1.76	.70	.7	-.5	.6	-.5
81	1010 0 0 0	2	7	-1.65	.92	.7	-.5	.6	-.4
84	1111 0	4	5	.69	1.18	.6	-.5	.4	-.4
138	1101000000010000	4	16	-1.33	.63	.8	-.5	.8	-.2
149	1101101101001000	8	16	-.01	.55	.9	-.5	.8	-.6
179	1111111111011010	13	16	1.76	.70	.7	-.5	.6	-.5
194	1111111011010010	11	16	.95	.59	.8	-.5	.7	-.6
197	1111100001011000	8	16	-.01	.55	.9	-.5	.8	-.6
221	1100000010000000	3	16	-1.76	.69	.8	-.5	.7	-.3
45	1111111010110010	11	16	.95	.59	.9	-.4	.7	-.5
61	1110101010101000	8	16	-.01	.55	.9	-.4	.8	-.5
72	1101110000000010	6	16	-.62	.57	.9	-.4	1.0	.0
91	1100 00 1 0 0	3	9	-1.22	.77	.8	-.4	.8	-.3
148	11 11 0	4	5	.82	1.17	.6	-.4	.4	-.3
158	1110111111100010	11	16	.95	.59	.9	-.4	.8	-.3
168	1100010110000000	5	16	-.96	.59	.9	-.4	.8	-.5
170	1111111110011100	12	16	1.32	.63	.8	-.4	.7	-.5
205	1110010101001000	7	16	-.31	.55	.9	-.4	.8	-.5
212	1011100000100000	5	16	-.96	.59	.9	-.4	.8	-.4
225	1111111100101100	11	16	.95	.59	.9	-.4	.7	-.5

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11					MNSQ	INFIT	MNSQ	OUTFIT
	3476218140653952	COUNT	TEST	MEASURE	ERROR				
226	1111111011111100	13	16	1.76	.70	.8	-.4	.7	-.2
241	1111100100111000	9	16	.30	.56	.9	-.4	.8	-.6
243	0111111110110000	10	16	.61	.57	.9	-.4	1.6	1.6
255	1111110110010100	10	16	.61	.57	.9	-.4	.8	-.5
4	10000000 0 0 0	1	11	-3.02	1.09	.6	-.3	.3	-.4
53	1111101111001010	11	16	.95	.59	.9	-.3	.8	-.4
95	10001000 0 0 0 0	2	12	-2.14	.83	.8	-.3	.6	-.1
120	1001 00 0 0	2	8	-1.81	.88	.8	-.3	.7	-.2
123	1101100000000010	5	16	-.96	.59	.9	-.3	1.1	.3
125	1110111110100010	10	16	.61	.57	.9	-.3	.8	-.4
139	1100111111001000	9	16	.30	.56	.9	-.3	.9	-.3
147	10 01010 0000	3	11	-.79	.71	.9	-.3	.7	-.3
165	1111110110110100	11	16	.95	.59	.9	-.3	.8	-.4
181	1111011100010100	9	16	.30	.56	.9	-.3	.8	-.4
257	1111111100111100	12	16	1.32	.63	.9	-.3	.7	-.4
2	1000001000000000	2	16	-2.31	.80	.8	-.2	.6	-.1
17	1010010010000000	4	16	-1.33	.63	.9	-.2	.8	-.3
34	11111111 1 1 0	10	11	2.17	1.08	.7	-.2	.3	-.4
52	1111111111111110	15	16	3.18	1.08	.6	-.2	.2	-.3
93	11 1 0 1 0	4	6	.42	.97	.8	-.2	.7	-.2
124	1110001110110000	8	16	-.01	.55	.9	-.2	.8	-.4
128	1101101101011000	9	16	.30	.56	.9	-.2	.9	-.3
173	1011101111111000	11	16	.95	.59	.9	-.2	1.2	.5
232	1111111111111110	15	16	3.18	1.08	.6	-.2	.2	-.3
31	1111111011111110	14	16	2.32	.81	.9	-.1	.9	.2
49	1111111110111110	14	16	2.32	.81	.8	-.1	.7	.0
54	1111110101111100	12	16	1.32	.63	.9	-.1	.8	-.1
64	1100010100001000	5	16	-.96	.59	.9	-.1	.9	-.1
82	1010 1 0 0	3	7	-.93	.84	.9	-.1	.8	-.3
89	01 10 0 0	2	6	-.87	.90	.9	-.1	.8	-.2
132	1001000001000000	3	16	-1.76	.69	.9	-.1	.8	-.1
156	1001010100000000	4	16	-1.33	.63	.9	-.1	.7	-.3
192	1111111110111110	14	16	2.32	.81	.8	-.1	.7	.0
213	1000100001000000	3	16	-1.76	.69	.9	-.1	.8	.0
261	1011010100001000	6	16	-.62	.57	1.0	-.1	.9	-.2
29	1111111011110110	13	16	1.76	.70	.9	.0	.8	-.1
62	1100000000101000	4	16	-1.33	.63	1.0	.0	1.1	.3
107	1000001001000000	3	16	-1.76	.69	1.0	.0	.9	.1
130	1100 0100000100	4	14	-1.06	.65	1.0	.0	1.2	.6
135	1010010011000000	5	16	-.96	.59	1.0	.0	.8	-.2
151	1111111101010110	12	16	1.32	.63	1.0	.0	.8	-.3
153	1111011001001010	9	16	.30	.56	1.0	.0	.9	-.2
203	1100001001001000	5	16	-.96	.59	1.0	.0	.9	.0
217	1001111100001000	7	16	-.31	.55	1.0	.0	.9	-.2
238	1111111011100110	12	16	1.32	.63	1.0	.0	.8	-.2

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11								
	3476218140653952	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
7	10000001010 0000	3	15	-1.71	.70	1.0	.1	1.0	.2
27	1000000000000001	2	16	-2.31	.80	.9	.1	6.6	3.0
76	1001000000000010	3	16	-1.76	.69	1.0	.1	1.7	1.0
106	1000110110000000	5	16	-.96	.59	1.0	.1	.8	-.2
113	01 0 0 0	1	5	-1.63	1.15	.9	.1	.7	.0
136	0111001000000000	4	16	-1.33	.63	1.0	.1	.8	-.3
155	1110110001010100	8	16	-.01	.55	1.0	.1	.9	-.1
164	1001000101000000	4	16	-1.33	.63	1.0	.1	.9	-.1
207	1111101111000110	11	16	.95	.59	1.0	.1	.9	-.2
248	1101100000100100	6	16	-.62	.57	1.0	.1	1.1	.3
250	0111110000010000	6	16	-.62	.57	1.0	.1	1.0	.2
252	1100101101110000	8	16	-.01	.55	1.0	.1	.9	-.3
254	1010011010100000	6	16	-.62	.57	1.0	.1	.9	-.2
25	1010011000000100	5	16	-.96	.59	1.0	.2	1.2	.5
46	1111111010110110	12	16	1.32	.63	1.0	.2	.9	-.1
56	0110010000000000	3	16	-1.76	.69	1.0	.2	.7	-.2
71	1000001000000010	3	16	-1.76	.69	1.0	.2	1.7	1.1
73	1010011001010000	6	16	-.62	.57	1.0	.2	.9	-.1
96	01 0	1	3	-1.30	1.23	1.0	.2	1.0	.1
126	1011001000101000	6	16	-.62	.57	1.0	.2	.9	.0
157	1011100001000100	6	16	-.62	.57	1.0	.2	1.1	.3
174	01111010 0 0 0	5	11	-.61	.65	1.0	.2	1.1	.5
193	1111110101111110	13	16	1.76	.70	1.0	.2	1.1	.3
206	1110110000110100	8	16	-.01	.55	1.0	.2	1.0	.0
235	1101110111111010	12	16	1.32	.63	1.0	.2	1.2	.5
16	0100000000000000	1	16	-3.15	1.08	1.0	.3	.5	.1
28	1111111001011110	12	16	1.32	.63	1.1	.3	.9	.1
38	1111111111110101	14	16	2.32	.81	1.1	.3	.7	-.1
42	1111 00110111010	10	15	.88	.60	1.1	.3	1.0	.1
47	1111111111111011	15	16	3.18	1.08	1.0	.3	.5	.1
63	1000010100100000	4	16	-1.33	.63	1.1	.3	1.0	.1
75	1011010001000010	6	16	-.62	.57	1.1	.3	1.1	.4
121	1101011110110100	10	16	.61	.57	1.1	.3	1.0	.2
210	1110101000110100	8	16	-.01	.55	1.0	.3	1.0	.0
222	1111110110010110	11	16	.95	.59	1.1	.3	.9	-.1
6	0010000000000000	1	16	-3.15	1.08	1.1	.4	.7	.3
41	1110101110111010	11	16	.95	.59	1.1	.4	1.0	.2
57	1101000111101000	8	16	-.01	.55	1.1	.4	1.0	.0
66	1010000001000001	4	16	-1.33	.63	1.1	.4	3.0	2.5
70	0101101001000000	5	16	-.96	.59	1.1	.4	1.0	.1
159	1111010001101010	9	16	.30	.56	1.1	.4	1.0	.0
187	1110111110110001	11	16	.95	.59	1.1	.4	1.1	.3
202	1111101101101001	11	16	.95	.59	1.1	.4	1.0	.2
15	0000000001000000	1	16	-3.15	1.08	1.2	.5	1.7	.9
19	1000001010001000	4	16	-1.33	.63	1.1	.5	1.1	.4

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11								
	3476218140653952	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
22	0000000001000000	1	16	-3.15	1.08	1.2	.5	1.7	.9
60	1001010100000010	5	16	-.96	.59	1.1	.5	1.2	.7
83	01000100 0 0 0	2	11	-2.13	.83	1.2	.5	1.0	.2
87	01 00 0	1	5	-2.42	1.20	1.2	.5	.9	.2
112	0000100 0 0 0	1	10	-2.95	1.10	1.2	.5	1.3	.6
211	1001000011100000	5	16	-.96	.59	1.1	.5	1.0	.1
216	1000001000101000	4	16	-1.33	.63	1.1	.5	1.2	.5
12	1000001000100010	4	16	-1.33	.63	1.2	.6	1.6	1.1
14	00000000 0 1 0	1	11	-3.02	1.09	1.3	.6	3.0	1.5
65	1011110000100001	7	16	-.31	.55	1.1	.6	1.6	1.7
67	1010 0000001100	4	14	-1.06	.65	1.2	.6	1.5	1.0
92	11010001 1 1 0	6	11	-.19	.65	1.1	.6	1.0	.2
109	0000 0001000000	1	14	-3.00	1.09	1.3	.6	1.7	.9
111	0010000100000000	2	16	-2.31	.80	1.2	.6	1.0	.3
114	0001 00 0 0 0	1	9	-2.85	1.12	1.3	.6	1.2	.6
137	1100100011000001	6	16	-.62	.57	1.1	.6	1.9	2.1
175	1111011010111110	12	16	1.32	.63	1.2	.6	1.1	.4
176	1111011111111111	15	16	3.18	1.08	1.3	.6	2.6	1.2
220	1100101001100010	7	16	-.31	.55	1.1	.6	1.1	.3
230	1111110100011110	11	16	.95	.59	1.1	.6	1.0	.1
231	1111111011100011	12	16	1.32	.63	1.2	.6	1.0	.1
259	1011111101010110	11	16	.95	.59	1.1	.6	1.3	.7
9	0010000000000001	2	16	-2.31	.80	1.3	.7	6.8	3.0
20	0100000100001000	3	16	-1.76	.69	1.3	.7	1.3	.7
24	0100 0 00000010	2	13	-2.03	.84	1.3	.7	2.4	1.3
36	1111100111011001	11	16	.95	.59	1.2	.7	1.1	.4
108	1100 0000000011	4	14	-1.06	.65	1.2	.7	3.2	2.8
190	1110111100001110	10	16	.61	.57	1.1	.7	1.0	.2
196	1011010010111000	8	16	-.01	.55	1.1	.7	1.0	.3
244	0110001001100000	5	16	-.96	.59	1.2	.7	1.1	.3
5	0000000001100000	2	16	-2.31	.80	1.4	.8	1.7	.9
98	11000001 1 1 0	5	11	-.61	.65	1.1	.8	1.1	.4
118	00001001 0 0 0	2	11	-2.13	.83	1.4	.8	1.4	.7
129	1000001011001000	5	16	-.96	.59	1.2	.8	1.1	.4
177	1111111101011111	14	16	2.32	.81	1.4	.8	1.2	.6
218	1001010101000010	6	16	-.62	.57	1.2	.8	1.2	.6
21	1001001001011000	6	16	-.62	.57	1.2	.9	1.1	.4
35	1110111000111001	10	16	.61	.57	1.2	.9	1.2	.7
117	0000 1000010000	2	14	-2.12	.83	1.4	.9	1.5	.8
133	1000001010101000	5	16	-.96	.59	1.2	.9	1.2	.5
134	0000011010000000	3	16	-1.76	.69	1.4	.9	1.2	.4
142	1000000001001010	4	16	-1.33	.63	1.3	.9	1.8	1.3
200	0000000010001000	2	16	-2.31	.80	1.4	.9	2.0	1.1
214	1100001100010001	6	16	-.62	.57	1.2	.9	2.0	2.2
253	1101010010001001	7	16	-.31	.55	1.2	.9	1.7	1.9

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1111 1 11								
	3476218140653952	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
68	1000 0011000001	4	14	-1.06	.65	1.3	1.0	2.9	2.5
26	000010000010100	3	15	-1.75	.69	1.5	1.1	1.7	1.2
88	0100010010000010	4	16	-1.33	.63	1.3	1.1	1.7	1.2
116	0001001101000000	4	16	-1.33	.63	1.3	1.1	1.1	.4
189	1111111001011111	13	16	1.76	.70	1.4	1.1	1.3	.6
201	0001011110000000	5	16	-.96	.59	1.3	1.1	1.1	.4
258	0010100010001000	4	16	-1.33	.63	1.3	1.1	1.3	.7
1	0100000110000001	4	16	-1.33	.63	1.4	1.2	3.3	2.8
141	1111101110010111	12	16	1.32	.63	1.4	1.2	1.2	.6
166	1000000101100100	5	16	-.96	.59	1.3	1.2	1.5	1.1
199	0100010000100001	4	16	-1.33	.63	1.4	1.2	3.3	2.8
23	0000100000010001	3	16	-1.76	.69	1.5	1.3	4. 4.7	3.0
58	0100001000001001	4	16	-1.33	.63	1.4	1.3	3.4	2.9
195	1110 00011100110	8	15	.19	.57	1.3	1.3	1.2	.7
228	0001010100000100	4	16	-1.33	.63	1.4	1.3	1.7	1.2
40	1111110001110111	12	16	1.32	.63	1.5	1.4	1.3	.7
227	1101100111000101	9	16	.30	.56	1.3	1.4	1.5	1.4
8	0000001011010000	4	16	-1.33	.63	1.5	1.5	1.4	.9
215	0000001011001000	4	16	-1.33	.63	1.5	1.5	1.5	1.0
246	0010000111001000	5	16	-.96	.59	1.4	1.6	1.4	.9
160	0101011100000110	7	16	-.31	.55	1.4	1.8	1.7	1.8
204	0000011110000100	5	16	-.96	.59	1.5	1.8	1.6	1.4
3	1000010001100110	6	16	-.62	.57	1.5	1.9	1.7	1.6
18	0000010000100011	4	16	-1.33	.63	1.7	2.0	4.1	3.4
131	1000011011011100	8	16	-.01	.55	1.4	2.1	1.4	1.2
74	0010110011001100	7	16	-.31	.55	1.5	2.2	1.7	1.8
110	1000100010101011	7	16	-.31	.55	1.6	2.8	2.2	2.9
48	1111111111111111	16	16	4.26	ESTIMATED FROM MAX MEASURE OBSERVED				
185	1111111111111111	16	16	4.26	ESTIMATED FROM MAX MEASURE OBSERVED				
262	11 111 1	6	6	4.26	ESTIMATED FROM MAX MEASURE OBSERVED				
79	00 000 0	0	6	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
99	0000 0 0	0	6	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
100	00 000 0	0	6	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
101	00 0 0	0	4	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
103	00	0	2	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
104	00	0	2	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
105	0	0	1	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
115	0	0	1	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
119	0000000 0 0	0	9	-4.23	ESTIMATED FROM MIN MEASURE OBSERVED				
	3476218111151911								
	1406 3 52								

ITEM STATISTICS -- INFIT ORDER

NUM	NAME	COUNT	SAMPLE	CALIBRTN	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
10	CAL62	103	235	.10	.15	1.3	3.9	1.5	4.2
13	EMP92	78	216	.62	.16	1.2	2.8	1.3	2.2
12	EMP91	28	216	2.32	.22	1.3	1.9	3.1	4.3
14	EMP93	100	216	.06	.16	1.1	1.2	1.1	.9
15	EMP94	56	216	1.25	.18	1.1	.9	1.3	1.7
16	EMP96	91	216	.29	.16	1.1	.9	1.1	.5
11	CAL69	111	230	-.11	.15	1.1	.8	1.1	.7
1	MAE17	120	225	-.37	.16	1.0	-.3	.9	-1.0
8	CAL45	123	241	-.31	.15	1.0	-.6	.9	-.8
9	CAL56	59	237	1.24	.17	.9	-.6	.9	-.7
3	NOC10	201	244	-2.28	.19	.8	-1.6	.6	-1.7
6	CAL02	138	245	-.61	.15	.9	-1.6	.8	-1.8
4	NOC15	165	244	-1.25	.16	.8	-2.3	.7	-2.2
7	CAL35	147	241	-.86	.15	.8	-2.8	.7	-2.3
5	NOC16	90	238	.43	.16	.8	-3.0	.8	-1.3
2	MAE18	124	222	-.51	.16	.8	-3.1	.9	-.8

MEAP MATH; PARTIAL CREDIT OF SHORT VERSION GRADE 10 "BIGSCALE" RASCH ANALYSIS VER. 1.73 Jan 31 20:28:18 1991
INPUT: 262 PERSONS 16 ITEMS ANALYZED: 258 PERSONS 16 ITEMS 4 CATEGORIES TABLE C-1

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1 3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
85		0	0	OUT					
103	22	4	2	.12	.71	.0	-3.2	.0	-3.5
119	22211 21 2 1	14	9	-.11	.32	.3	-3.2	.2	-2.9
115	1	1	1	-.76	1.00	.0	-2.4	.0	-1.8
99	1211 1 1	7	6	-.61	.41	.3	-2.3	.3	-2.0
105	2	2	1	.15	1.00	.0	-2.3	.0	-3.0
69	3323 22 11223222	30	14	.67	.29	.4	-2.0	.4	-1.7
219	33331111221221222	30	16	.36	.25	.5	-2.0	.6	-1.7
164	31232221332210l0	28	16	.24	.25	.6	-1.8	.6	-1.7
13	31 1 1	6	4	-.34	.49	.3	-1.7	.3	-1.6
208	33333121222011120	29	16	.30	.25	.6	-1.7	.6	-1.6
79	11 21 2 1	8	6	-.41	.40	.4	-1.6	.5	-1.5
123	33233222121222300	31	16	.43	.25	.6	-1.5	.6	-1.5
127	33323231323212222	37	16	.87	.29	.5	-1.5	.5	-1.2
169	33331210222212200	27	16	.18	.24	.7	-1.5	.7	-1.4
81	3231 2 1 0	12	7	.02	.38	.5	-1.4	.5	-1.3
168	33212313132122222	33	16	.56	.26	.6	-1.4	.6	-1.1
172	33332233233212l2	38	16	.96	.31	.5	-1.4	.5	-1.3
155	33323223322231l2	39	16	1.05	.32	.5	-1.3	.5	-1.2
157	32333112322221l2	35	16	.70	.28	.6	-1.3	.6	-1.1
221	33211302012112l1	25	16	.06	.24	.7	-1.3	.7	-1.3
82	3232 3 1 1	15	7	.44	.41	.5	-1.2	.5	-1.1
89	23 3 11 1	11	6	.28	.40	.5	-1.2	.5	-1.0
95	32023 1221 2 01	19	12	.03	.28	.7	-1.2	.7	-1.1
120	3223 0 11 0	12	8	-.18	.34	.6	-1.2	.6	-1.4
171	33333133223212l2	37	16	.87	.29	.6	-1.1	.6	-1.0
220	33223232323123l0	36	16	.78	.28	.6	-1.1	.6	-1.2
243	23333333233232222	42	16	1.42	.39	.5	-1.1	.7	-.4
248	33233211123222l2	35	16	.70	.28	.6	-1.1	.6	-1.0
71	32212132210123l2	29	16	.39	.25	.7	-1.0	.7	-1.1
75	32331213312213l2	33	16	.56	.26	.7	-1.0	.7	-.9
76	32132202221213l2	28	16	.24	.25	.7	-1.0	.7	-1.0
91	3321 1 31 1 0	15	9	.03	.33	.7	-1.0	.7	-1.0
244	13322231323112l2	33	16	.56	.26	.7	-1.0	1.0	.1
96	23 2	7	3	.61	.67	.3	-.9	.3	-.9
126	32332230213312l1	33	16	.56	.26	.7	-.9	.7	-.8
147	31 12 3312 212	21	11	.50	.30	.7	-.9	.8	-.7
166	32222221333220l0	32	16	.49	.26	.7	-.9	.7	-1.1
200	21122312010310l1	22	16	-.12	.24	.8	-.9	.8	-.8
225	33333233233322l2	43	16	1.59	.42	.5	-.9	.5	-.8
241	33333212233332l2	40	16	1.16	.34	.6	-.9	.6	-.6
14	12211 1021 3 1	15	11	-.25	.29	.8	-.8	.8	-.5
32	33333331332232l2	41	16	1.28	.36	.6	-.8	.6	-.6

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1 3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ INFIT		MNSQ OUTFIT	
56	1332222321001020	24	16	.00	.24	.8	-.8	.9	-.5
62	3312121120331122	28	16	.24	.25	.8	-.8	.8	-.7
94	3331222313012020	28	16	.24	.25	.8	-.8	.8	-.8
97	33 22 0 1	11	6	.06	.41	.7	-.8	.6	-.8
100	20 12 0 1	6	6	-.75	.42	.6	-.8	.6	-.9
143	3331321321331210	32	16	.49	.26	.8	-.8	.8	-.7
149	3313323233131221	36	16	.78	.28	.7	-.8	.8	-.5
183	3333332323223210	38	16	.96	.31	.7	-.8	.5	-1.1
247	3333333333323222	44	16	1.79	.47	.5	-.8	.4	-.9
261	3233211323132012	32	16	.49	.26	.8	-.8	.7	-.8
55	3331220212031122	28	16	.24	.25	.8	-.7	.8	-.7
83	13212 1311 0 2	17	11	-.08	.29	.8	-.7	.9	-.4
87	23 1 2 0	8	5	-.15	.44	.7	-.7	.6	-.8
125	3332333323321322	41	16	1.28	.36	.6	-.7	.6	-.7
129	3122233130231210	29	16	.30	.25	.8	-.7	.8	-.8
181	3333223313223131	38	16	.96	.31	.7	-.7	.6	-.8
212	3133322102321102	29	16	.30	.25	.8	-.7	.8	-.7
224	3333333321311210	35	16	.70	.28	.8	-.7	.7	-.8
230	3333322323233332	43	16	1.59	.42	.6	-.7	.6	-.5
237	3333132331213210	34	16	.63	.27	.8	-.7	.7	-.7
4	30120 1200 1 0	10	11	-.71	.32	.8	-.6	.7	-.6
44	3333323333323232	44	16	1.79	.47	.5	-.6	.5	-.6
70	2323323031101011	26	16	.12	.24	.8	-.6	.8	-.7
116	2223023133220211	29	16	.30	.25	.8	-.6	.8	-.6
145	3332333333211220	37	16	.87	.29	.7	-.6	.7	-.8
154	3323323323021201	33	16	.56	.26	.8	-.6	.8	-.7
162	3333333232201212	36	16	.78	.28	.8	-.6	.7	-.7
215	1122133131131120	26	16	.12	.24	.8	-.6	.9	-.2
249	3331333323332222	41	16	1.28	.36	.7	-.6	.9	-.1
252	3321323133323222	38	16	.96	.31	.7	-.6	.9	-.1
259	3233323333223332	43	16	1.59	.42	.6	-.6	.7	-.4
66	3232121131021203	27	16	.18	.24	.9	-.5	.9	-.4
80	3331 1 32 1 0	17	9	.25	.34	.8	-.5	.8	-.5
102	32	5	2	.47	.95	.2	-.5	.2	-.3
130	3312 02 23211231	26	14	.36	.27	.8	-.5	.8	-.4
151	3333323333223332	44	16	1.79	.47	.6	-.5	.6	-.5
161	3333322103321220	33	16	.56	.26	.8	-.5	.8	-.5
175	3333233322333332	44	16	1.79	.47	.6	-.5	.6	-.4
233	3333231321031122	33	16	.56	.26	.8	-.5	.8	-.6
234	3333333333333222	45	16	2.04	.55	.5	-.5	.4	-.8
16	1301022120101010	15	16	-.56	.26	.8	-.4	.8	-.6
33	3333333132123011	35	16	.70	.28	.8	-.4	.8	-.5
43	3333333333323232	45	16	2.04	.55	.6	-.4	.5	-.6
51	3333333333233322	45	16	2.04	.55	.6	-.4	.5	-.6
90	33123 2323 1 0	23	11	.46	.32	.8	-.4	.8	-.4

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1 3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
163	3333332123230220	35	16	.70	.28	.8	-.4	.8	-.6
167	3333333323323321	43	16	1.59	.42	.7	-.4	.5	-.7
184	3333323333213232	42	16	1.42	.39	.7	-.4	.7	-.4
211	3223230230321220	30	16	.36	.25	.9	-.4	.8	-.5
242	3333333331321110	36	16	.78	.28	.8	-.4	.7	-.6
86	33 1 0 1	8	5	-.15	.44	.8	-.3	.8	-.4
124	3332233023313211	35	16	.70	.28	.9	-.3	.9	-.3
131	3122233331233232	38	16	.96	.31	.8	-.3	1.0	.1
153	3333213331231322	38	16	.96	.31	.8	-.3	.8	-.3
165	3333331323323232	42	16	1.42	.39	.8	-.3	.8	-.2
197	3333311131133122	34	16	.63	.27	.9	-.3	.9	-.2
198	3331012223011202	26	16	.12	.24	.9	-.3	.9	-.2
205	3332220333132220	34	16	.63	.27	.9	-.3	.9	-.3
206	3331321321323232	37	16	.87	.29	.8	-.3	.9	-.1
37	3333333323323132	43	16	1.59	.42	.8	-.2	.6	-.4
72	3313312312012322	32	16	.49	.26	.9	-.2	.9	-.1
78	33312 3313 1 2	25	11	.68	.34	.8	-.2	.9	-.2
93	33 3 2 3 1	15	6	1.03	.55	.7	-.2	.6	-.3
122	3333233322303222	39	16	1.05	.32	.9	-.2	.8	-.3
186	3333323333333212	43	16	1.59	.42	.8	-.2	.6	-.5
191	3333333123323210	38	16	.96	.31	.9	-.2	.8	-.5
193	3333322333333332	45	16	2.04	.55	.7	-.2	.7	-.2
204	1212133313211231	30	16	.36	.25	.9	-.2	1.1	.4
11	30331 22 1	15	8	.11	.35	.9	-.1	.9	.0
49	3333333323333332	46	16	2.42	.68	.7	-.1	.6	-.2
53	3333333133231322	41	16	1.28	.36	.9	-.1	.8	-.1
77	32	5	2	.76	.94	.4	-.1	.4	-.2
104	01	1	2	-1.42	.94	.5	-.1	.5	-.4
187	3331333323323223	42	16	1.42	.39	.9	-.1	1.1	.3
196	3133131321333222	36	16	.78	.28	.9	-.1	1.0	.2
239	3333323023331221	37	16	.87	.29	.9	-.1	.9	.0
260	33313 32 2	20	8	.90	.47	.8	-.1	.9	.1
46	3333333321323332	43	16	1.59	.42	.9	.0	.8	.0
57	3323131033331111	32	16	.49	.26	1.0	.0	1.0	.0
65	3133322320321223	35	16	.70	.28	1.0	.0	1.0	.1
112	10223 112 0 2	14	10	-.24	.31	1.0	.0	1.0	.1
146	3333 23 23301231	32	14	.85	.31	1.0	.0	.9	-.2
148	33 3 3 2	14	5	1.75	.95	.5	.0	.4	-.2
202	3333323233331123	41	16	1.28	.36	.9	.0	.8	-.3
250	0333321322223222	35	16	.70	.28	1.0	.0	2.0	2.2
6	2032202100000000	12	16	-.79	.29	1.0	.1	.9	-.2
7	322200203321 001	21	15	-.11	.25	1.0	.1	1.0	.0
12	3221003111321310	24	16	.00	.24	1.0	.1	1.0	.0
135	3231131331000212	26	16	.12	.24	1.0	.1	1.0	.0
140	3330322333223220	36	16	.78	.28	1.0	.1	1.1	.4

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1 3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ INFIT		MNSQ OUTFIT	
177	3333323333233333	46	16	2.42	.68	.9	.1	.9	.2
182	3333323333221032	39	16	1.05	.32	1.0	.1	.8	-.3
223	3333333333223202	41	16	1.28	.36	1.0	.1	.8	-.3
236	3330132331210212	30	16	.36	.25	1.0	.1	1.0	.2
254	3230233321320222	33	16	.56	.26	1.0	.1	1.1	.3
255	3333331313223230	38	16	.96	.31	1.0	.1	.9	-.2
256	3333333302021212	34	16	.63	.27	1.0	.1	1.0	.0
15	1011100030200111	12	16	-.79	.29	1.0	.2	1.1	.4
34	33333 3333 3 2	32	11	2.58	.97	.7	.2	.5	-.1
84	3333 1	13	5	1.03	.67	.9	.2	.6	-.1
101	00 1 1	2	4	-1.53	.67	.9	.2	1.1	.4
128	3323323133033222	38	16	.96	.31	1.0	.2	1.0	.2
133	3022233210332220	30	16	.36	.25	1.0	.2	1.1	.3
159	3333221331331302	36	16	.78	.28	1.0	.2	.9	-.1
190	3332323313231330	38	16	.96	.31	1.0	.2	.9	-.2
201	2013133323111021	27	16	.18	.24	1.0	.2	1.0	.2
213	3021321130001122	22	16	-.12	.24	1.0	.2	1.0	.2
216	3021023222331011	26	16	.12	.24	1.0	.2	1.0	.1
245	33 3333303321210	33	15	.74	.29	1.0	.2	1.0	.1
17	3032031321122002	25	16	.06	.24	1.1	.3	1.1	.3
41	3332333123333320	40	16	1.16	.34	1.1	.3	.9	.0
138	3313200102003111	21	16	-.18	.24	1.1	.3	1.0	.2
158	3332333333322302	41	16	1.28	.36	1.1	.3	.9	-.1
170	3333333323233230	42	16	1.42	.39	1.1	.3	.7	-.4
174	13333 3101 2 1	21	11	.27	.30	1.1	.3	1.3	.8
176	3333233333333333	47	16	3.08	.98	.9	.3	1.2	.6
209	3233323331001202	31	16	.43	.25	1.0	.3	1.0	.0
210	3332323101313132	34	16	.63	.27	1.1	.3	1.0	.1
251	3333333303321210	36	16	.78	.28	1.1	.3	1.0	.1
19	3011133101231200	22	16	-.12	.24	1.1	.4	1.1	.3
59	3333331023320222	35	16	.70	.28	1.1	.4	1.1	.3
194	3333333332023321	40	16	1.16	.34	1.1	.4	1.0	.2
207	3333333233220332	41	16	1.28	.36	1.1	.4	1.0	.1
64	3320021323132222	31	16	.43	.25	1.1	.5	1.2	.6
150	3333133332203012	35	16	.70	.28	1.1	.5	1.0	.2
152	3333 33 3321 302	32	13	1.10	.36	1.2	.5	1.0	.1
160	1323123323220332	35	16	.70	.28	1.1	.5	1.5	1.2
189	3333323331233333	44	16	1.79	.47	1.2	.5	1.1	.4
35	3332323320333213	39	16	1.05	.32	1.2	.6	1.1	.3
38	3333333333323133	45	16	2.04	.55	1.3	.6	.9	.2
60	3213012313201302	27	16	.18	.24	1.1	.6	1.1	.5
114	0213 220 1 0	11	9	-.41	.33	1.2	.6	1.2	.6
117	2122 03 20103022	20	14	-.05	.26	1.1	.6	1.1	.6
134	1111033312201222	25	16	.06	.24	1.1	.6	1.2	.9
141	3333333113223333	42	16	1.42	.39	1.2	.6	1.2	.5

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1							
	3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ OUTFIT
142	3201211032231312	27	16	.18	.24	1.1	.6	1.1 .5
144	3333333330233210	38	16	.96	.31	1.2	.6	1.0 .3
235	3313331333333312	41	16	1.28	.36	1.2	.6	1.4 .9
257	3333313323333230	41	16	1.28	.36	1.2	.6	1.0 .2
1	1322132013012013	25	16	.06	.24	1.2	.7	1.3 1.1
2	3002003201220010	16	16	-.49	.26	1.2	.7	1.1 .4
10	33032 00 1	12	8	-.24	.34	1.2	.7	1.2 .6
26	102232011133202	23	15	-.01	.25	1.2	.7	1.2 .9
30	3333333323333230	43	16	1.59	.42	1.3	.7	.8 -.1
121	3303233313323231	38	16	.96	.31	1.2	.7	1.5 1.2
139	3301333333231120	34	16	.63	.27	1.2	.7	1.4 1.0
179	3333333333233310	42	16	1.42	.39	1.3	.7	.9 -.1
180	3333333133333022	41	16	1.28	.36	1.3	.7	1.1 .4
229	333332010332 102	29	15	.40	.26	1.2	.7	1.1 .5
28	3333323332033331	41	16	1.28	.36	1.4	.8	1.2 .6
36	3333331233033223	40	16	1.16	.34	1.4	.8	1.3 .7
107	3020203232021020	22	16	-.12	.24	1.2	.8	1.2 .6
173	3033333233333122	40	16	1.16	.34	1.3	.8	1.8 1.4
228	2023012313220130	25	16	.06	.24	1.2	.8	1.2 .7
25	3032203322021030	26	16	.12	.24	1.2	.9	1.2 .8
29	3333333331323330	42	16	1.42	.39	1.4	.9	1.0 .2
47	3333333333333313	46	16	2.42	.68	1.7	.9	1.1 .5
67	3030 22 21131232	25	14	.29	.27	1.2	.9	1.3 1.0
106	3011332303202100	24	16	.00	.24	1.2	.9	1.2 .8
217	3003323323231011	30	16	.36	.25	1.2	.9	1.4 1.2
238	3333333330321332	41	16	1.28	.36	1.4	.9	1.3 .7
22	00010211322010l0	14	16	-.63	.27	1.3	1.0	1.2 .7
24	2301 20 0 020310	14	13	-.41	.29	1.3	1.0	1.3 1.0
27	3021000102200013	15	16	-.56	.26	1.3	1.0	1.7 1.9
42	3333 31023333320	35	15	.91	.31	1.4	1.0	1.3 .8
50	3333333333333230	44	16	1.79	.47	1.6	1.0	.9 .0
54	3333322333333030	40	16	1.16	.34	1.4	1.0	1.0 .2
111	0031122103212020	20	16	-.24	.25	1.2	1.0	1.3 1.1
132	3203200230001022	20	16	-.24	.25	1.3	1.0	1.3 1.1
136	0333223001222022	27	16	.18	.24	1.2	1.0	1.6 1.9
156	3023210303021002	22	16	-.12	.24	1.2	1.0	1.2 1.0
178	3333323333303230	40	16	1.16	.34	1.5	1.0	1.2 .5
214	3312223113003023	29	16	.30	.25	1.3	1.0	1.2 .8
218	3103220333111322	30	16	.36	.25	1.3	1.0	1.3 1.0
222	3333331303223330	38	16	.96	.31	1.4	1.0	1.3 .7
226	3333333330333232	43	16	1.59	.42	1.6	1.0	1.5 .9
45	3333333301323302	38	16	.96	.31	1.4	1.1	1.3 .8
61	3330333101332200	30	16	.36	.25	1.3	1.1	1.3 1.0
73	3232103331003022	28	16	.24	.25	1.3	1.1	1.2 .8
92	33230 1133 3 0	22	11	.37	.31	1.4	1.1	1.3 .8

OBSERVATIONS SORTED BY INFIT:
PERSON/ITEM

	1 1111 1 1								
	3476248101635592	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
113	03 2 0 2	7	5	-.14	.43	1.5	1.1	1.5	1.1
192	3333333323333330	44	16	1.79	.47	1.7	1.1	1.0	.2
246	1230230233230011	26	16	.12	.24	1.3	1.1	1.3	1.3
31	3333333331333330	43	16	1.59	.42	1.7	1.2	1.2	.5
39	3333333323333030	41	16	1.28	.36	1.6	1.2	1.1	.4
40	3333321330323333	41	16	1.28	.36	1.6	1.2	1.6	1.1
68	3121 30 31001023	20	14	-.05	.26	1.3	1.2	1.4	1.3
98	33220 1033 3 2	22	11	.37	.31	1.4	1.2	1.3	.9
109	0022 01 30210020	13	14	-.56	.29	1.4	1.2	1.4	1.1
118	10123 2003 1 2	15	11	-.25	.29	1.4	1.2	1.4	1.1
137	3310331230022113	28	16	.24	.25	1.3	1.2	1.3	1.1
258	1230331102032202	25	16	.06	.24	1.3	1.2	1.4	1.4
21	3003123231233012	29	16	.30	.25	1.4	1.3	1.4	1.4
74	2130332330131232	32	16	.49	.26	1.4	1.3	1.5	1.4
203	3301003231232102	26	16	.12	.24	1.3	1.3	1.3	1.3
9	0032221011021223	22	16	-.12	.24	1.4	1.4	1.6	2.1
240	3332303001003212	26	16	.12	.24	1.4	1.4	1.3	1.2
20	0311021021023232012	23	16	-.06	.24	1.4	1.5	1.6	2.0
52	3333333333333330	45	16	2.04	.55	2.2	1.5	1.1	.4
188	3333333103333002	36	16	.78	.28	1.5	1.5	1.4	1.0
231	3333333331301303	38	16	.96	.31	1.7	1.5	1.4	.9
232	3333333333333330	45	16	2.04	.55	2.2	1.5	1.1	.4
195	3330 30131311332	30	15	.51	.27	1.5	1.6	1.6	1.5
253	3303232320032203	31	16	.43	.25	1.5	1.6	1.5	1.4
8	10020331312231 00	22	16	-.12	.24	1.5	1.7	1.5	1.6
63	3001001323311212	23	16	-.06	.24	1.4	1.7	1.4	1.6
18	2002121311301323	25	16	.06	.24	1.5	1.8	1.6	1.9
23	2021301100003223	20	16	-.24	.25	1.6	1.9	1.8	2.3
88	0302130302222310	24	16	.00	.24	1.5	1.9	1.7	2.4
3	3121010330321332	28	16	.24	.25	1.6	2.0	1.5	1.7
227	3313330133002133	32	16	.49	.26	1.6	2.0	1.6	1.6
5	0002001130320022	16	16	-.49	.26	1.8	2.4	2.0	2.5
58	0320223002032213	25	16	.06	.24	1.7	2.4	2.0	3.0
108	3302 00 00200303	16	14	-.33	.27	1.8	2.4	2.2	2.8
110	3011330120332323	30	16	.36	.25	1.7	2.4	1.8	2.2
199	0322020300321223	25	16	.06	.24	1.7	2.4	2.0	3.0
48	3333333333333333	48	16	4.06	ESTIMATED FROM MAX MEASURE OBSERVED				
185	3333333333333333	48	16	4.06	ESTIMATED FROM MAX MEASURE OBSERVED				
262	33 33 3 3	18	6	4.06	ESTIMATED FROM MAX MEASURE OBSERVED				
	3476218111115191								
	4 0163 5 2								

ITEM STATISTICS -- INFIT ORDER

NUM	NAME	COUNT	SAMPLE	CALIBRTN	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
12	EMP91	311	216	.75	.07	1.4	4.8	1.6	5.7
3	NOC10	653	249	-.96	.10	1.3	2.2	1.0	-.2
4	NOC15	575	249	-.37	.08	1.2	2.0	1.0	.0
10	CAL62	476	235	.05	.07	1.2	2.0	1.1	1.2
11	CAL69	459	230	.11	.07	1.1	1.4	1.2	1.9
1	MAE17	476	228	-.02	.07	1.1	1.2	1.0	.4
16	EMP96	432	216	.13	.07	1.1	1.1	1.1	.6
8	CAL45	507	243	-.04	.07	1.0	.4	1.0	-.2
13	EMP92	422	216	.18	.07	1.0	.0	1.0	-.4
7	CAL35	566	245	-.36	.08	1.0	-.3	.8	-1.2
6	CAL02	572	254	-.30	.08	.9	-.7	.9	-.8
2	MAE18	496	225	-.17	.08	.9	-1.0	.8	-1.4
15	EMP94	378	216	.41	.07	.9	-1.3	.9	-.7
9	CAL56	402	238	.43	.07	.9	-1.7	.9	-.7
14	EMP93	475	216	-.12	.08	.8	-1.8	.8	-1.5
5	NOC16	439	243	.27	.07	.8	-3.3	.8	-2.7

MEAP Math 01-07-1991 20:51:11
Table D-1 Students Measurement Report (ordered by Infit).

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
415	415	48	16		Maximum						0.00	415
753	753	48	16		Maximum						0.00	753
890	890	18	6		Maximum						0.00	890
477	477	4	2	2.0	0.20	0.69	0.0	-3	0.0	-3	0.00	477
493	493	14	9	1.6	-0.23	0.31	0.2	-3	0.3	-3	0.22	493
479	479	2	1	2.0	0.01	1.04	0.0	-2	0.0	-2	0.00	479
452	452	8	6	1.3	-0.27	0.38	0.3	-2	0.3	-2	-0.07	452
380	380	6	4	1.5	-0.41	0.46	0.3	-2	0.3	-1	0.49	380
473	473	7	6	1.2	-0.45	0.39	0.3	-2	0.2	-2	0.07	473
816	816	30	16	1.9	0.20	0.25	0.5	-2	0.5	-2	0.24	816
489	489	1	1	1.0	-1.06	1.13	0.0	-1	0.0	-1	0.00	489
469	469	7	3	2.3	0.67	0.65	0.3	-1	0.3	0	0.09	469
711	711	38	16	2.4	0.80	0.31	0.5	-1	0.4	-1	0.43	711
437	437	30	14	2.1	0.72	0.30	0.5	-1	0.4	-1	0.24	437
665	665	37	16	2.3	0.71	0.30	0.5	-1	0.5	-1	0.33	665
462	462	11	6	1.8	0.35	0.40	0.5	-1	0.5	-1	0.43	462
455	455	15	7	2.1	0.27	0.40	0.5	-1	0.4	-1	0.49	455
454	454	12	7	1.7	-0.11	0.36	0.5	-1	0.6	-1	0.58	454
474	474	6	6	1.0	-0.57	0.41	0.5	-1	0.5	0	0.19	474
694	694	39	16	2.4	0.91	0.33	0.6	-1	0.5	0	0.24	694
696	696	35	16	2.2	0.75	0.28	0.6	-1	0.6	0	0.19	696
710	710	37	16	2.3	0.71	0.30	0.6	-1	0.6	0	0.40	710
661	661	31	16	1.9	0.47	0.25	0.6	-1	0.6	-1	0.46	661
707	707	33	16	2.1	0.40	0.26	0.6	-1	0.6	-1	0.17	707
805	805	29	16	1.8	0.35	0.24	0.6	-1	0.6	-1	0.49	805
703	703	28	16	1.8	0.29	0.24	0.6	-1	0.6	-1	0.45	703
849	849	25	16	1.6	0.12	0.24	0.6	-1	0.6	-1	0.26	849
468	468	19	12	1.6	0.10	0.27	0.6	-1	0.6	-1	0.32	468
708	708	27	16	1.7	0.02	0.24	0.6	-1	0.6	-1	0.48	708
494	494	12	8	1.5	-0.07	0.32	0.6	-1	0.7	-1	0.52	494
381	381	15	11	1.4	-0.14	0.28	0.6	-1	0.8	0	-0.12	381
872	872	33	16	2.1	0.61	0.26	0.7	-1	0.9	0	0.04	872
443	443	33	16	2.1	0.40	0.26	0.7	-1	0.6	-1	0.23	443
444	444	28	16	1.8	0.29	0.24	0.7	-1	0.7	-1	0.19	444
439	439	29	16	1.8	0.14	0.24	0.7	-1	0.7	-1	0.03	439
430	430	28	16	1.8	0.08	0.24	0.7	-1	0.7	-1	0.14	430
464	464	15	9	1.7	-0.10	0.32	0.7	-1	0.7	0	0.49	464
797	797	22	16	1.4	-0.26	0.24	0.7	-1	0.7	-1	0.18	797
476	476	5	2	2.5	0.26	0.77	0.1	0	0.1	0	0.37	476
686	686	14	5	2.8	1.43	0.87	0.3	0	0.2	0	0.51	686
475	475	2	4	0.5	-1.32	0.70	0.3	0	0.3	0	-0.44	475
478	478	1	2	0.5	-1.40	1.00	0.3	0	0.3	0	-0.08	478
862	862	45	16	2.8	2.03	0.60	0.4	0	0.3	0	0.50	862
875	875	44	16	2.8	1.73	0.51	0.4	0	0.3	0	0.49	875
450	450	5	2	2.5	0.61	0.91	0.4	0	0.4	0	0.12	450
401	401	32	11	2.9	2.55	0.96	0.5	0	0.2	0	0.37	401
410	410	45	16	2.8	2.24	0.60	0.5	0	0.4	0	0.40	410

MEAP Math 01-07-1991 20:51:11
Table D-1 Students Measurement Report (ordered by Infit) - cont'd

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
411	411	44	16	2.8	1.73	0.51	0.5	0	0.5	0	0.30	411
871	871	42	16	2.6	1.52	0.41	0.5	0	0.9	0	0.25	871
853	853	43	16	2.7	1.50	0.45	0.5	0	0.5	0	0.34	853
418	418	45	16	2.8	2.03	0.60	0.6	0	0.4	0	0.40	418
690	690	44	16	2.8	1.73	0.51	0.6	0	0.6	0	0.21	690
663	663	41	16	2.6	1.37	0.38	0.6	0	0.5	0	0.35	663
470	470	11	6	1.8	-0.05	0.39	0.6	0	0.6	-1	0.48	470
416	416	46	16	2.9	2.67	0.73	0.7	0	0.6	0	0.29	416
761	761	45	16	2.8	2.03	0.60	0.7	0	0.8	0	0.16	761
706	706	43	16	2.7	1.50	0.45	0.7	0	0.5	0	0.45	706
754	754	43	16	2.7	1.50	0.45	0.7	0	0.6	0	0.39	754
858	858	43	16	2.7	1.50	0.45	0.7	0	0.6	0	0.15	858
887	887	43	16	2.7	1.50	0.45	0.7	0	0.9	0	0.10	887
877	877	41	16	2.6	1.37	0.38	0.7	0	0.8	0	0.28	877
399	399	41	16	2.6	1.16	0.38	0.7	0	0.6	0	0.33	399
466	466	15	6	2.5	1.09	0.55	0.7	0	0.5	0	0.42	466
869	869	40	16	2.5	1.02	0.35	0.7	0	0.6	0	0.24	869
751	751	38	16	2.4	1.01	0.31	0.7	0	0.6	0	0.53	751
888	888	20	8	2.5	0.92	0.45	0.7	0	0.7	0	0.19	888
749	749	38	16	2.4	0.91	0.31	0.7	0	0.7	0	0.47	749
457	457	13	5	2.6	0.83	0.64	0.7	0	0.5	0	0.56	457
876	876	35	16	2.2	0.75	0.28	0.7	0	0.6	0	0.15	876
683	683	37	16	2.3	0.71	0.30	0.7	0	0.7	0	0.51	683
688	688	36	16	2.3	0.62	0.29	0.7	0	0.7	0	0.32	688
701	701	36	16	2.3	0.62	0.29	0.7	0	0.7	0	0.42	701
817	817	36	16	2.3	0.62	0.29	0.7	0	0.6	0	0.37	817
664	664	33	16	2.1	0.61	0.26	0.7	0	0.7	0	0.30	664
685	685	21	11	1.9	0.54	0.31	0.7	0	0.7	0	0.03	685
852	852	35	16	2.2	0.54	0.28	0.7	0	0.7	0	0.56	852
681	681	32	16	2.0	0.33	0.26	0.7	0	0.7	0	0.43	681
456	456	17	11	1.5	0.02	0.28	0.7	0	0.8	0	0.11	456
460	460	8	5	1.6	-0.22	0.41	0.7	0	0.7	0	0.45	460
369	369	10	11	0.9	-0.57	0.31	0.7	0	0.7	0	0.36	369
743	743	44	16	2.8	1.94	0.51	0.8	0	0.8	0	0.19	743
404	404	43	16	2.7	1.50	0.45	0.8	0	0.6	0	0.35	404
752	752	42	16	2.6	1.31	0.41	0.8	0	0.7	0	0.25	752
692	692	38	16	2.4	1.01	0.31	0.8	0	0.8	0	0.22	692
870	870	36	16	2.3	0.83	0.29	0.8	0	0.7	0	0.58	870
880	880	38	16	2.4	0.80	0.31	0.8	0	0.8	0	0.09	880
865	865	34	16	2.1	0.68	0.27	0.8	0	0.8	0	0.44	865
700	700	33	16	2.1	0.61	0.26	0.8	0	0.8	0	0.42	700
702	702	35	16	2.2	0.54	0.28	0.8	0	0.8	0	0.48	702
889	889	32	16	2.0	0.54	0.26	0.8	0	0.8	0	0.28	889
451	451	25	11	2.3	0.53	0.34	0.8	0	0.8	0	0.28	451
693	693	33	16	2.1	0.40	0.26	0.8	0	0.8	0	0.44	693
861	861	33	16	2.1	0.40	0.26	0.8	0	0.8	0	0.33	861
667	667	29	16	1.8	0.35	0.24	0.8	0	0.7	-1	0.34	667
705	705	32	16	2.0	0.33	0.26	0.8	0	0.7	0	0.30	705

MEAP Math 01-07-1991 20:51:11
Table D-1 Students Measurement Report (ordered by Infit) - cont'd

Num	Students	Score	Count	Average	Measure Logit	Measure Error	Model Infit MnSq	Model Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
453	453	17	9	1.9	0.32	0.33	0.8	0	0.8	0	0.45	453
463	463	23	11	2.1	0.32	0.31	0.8	0	0.8	0	0.41	463
423	423	28	16	1.8	0.29	0.24	0.8	0	0.7	0	0.21	423
668	668	26	14	1.9	0.20	0.27	0.8	0	0.9	0	0.13	668
808	808	30	16	1.9	0.20	0.25	0.8	0	0.8	0	0.35	808
438	438	26	16	1.6	0.18	0.24	0.8	0	0.8	0	0.36	438
490	490	29	16	1.8	0.14	0.24	0.8	0	0.8	0	0.20	490
809	809	29	16	1.8	0.14	0.24	0.8	0	0.8	0	0.33	809
467	467	28	16	1.8	0.08	0.24	0.8	0	0.8	0	0.41	467
424	424	24	16	1.5	0.06	0.24	0.8	0	0.8	0	0.34	424
434	434	27	16	1.7	0.02	0.24	0.8	0	0.8	0	0.16	434
378	378	15	8	1.9	-0.00	0.34	0.8	0	0.7	0	0.27	378
812	812	26	16	1.6	-0.03	0.24	0.8	0	0.8	0	0.14	812
459	459	8	5	1.6	-0.22	0.41	0.8	0	0.7	0	0.48	459
486	486	14	10	1.4	-0.34	0.29	0.8	0	0.8	0	-0.01	486
383	383	15	16	0.9	-0.46	0.25	0.8	0	0.7	-1	0.23	383
704	704	42	16	2.6	1.52	0.41	0.9	0	0.8	0	0.24	704
660	660	39	16	2.4	1.12	0.33	0.9	0	0.8	0	0.29	660
759	759	38	16	2.4	1.01	0.31	0.9	0	0.8	0	0.48	759
400	400	35	16	2.2	0.75	0.28	0.9	0	0.8	0	0.45	400
803	803	37	16	2.3	0.71	0.30	0.9	0	0.9	0	0.08	803
867	867	37	16	2.3	0.71	0.30	0.9	0	0.8	0	0.35	867
794	794	34	16	2.1	0.68	0.27	0.9	0	0.9	0	0.20	794
440	440	32	16	2.0	0.54	0.26	0.9	0	0.9	0	0.11	440
662	662	35	16	2.2	0.54	0.28	0.9	0	0.8	0	0.31	662
802	802	34	16	2.1	0.47	0.27	0.9	0	0.8	0	0.33	802
864	864	30	16	1.9	0.41	0.25	0.9	0	1.0	0	0.25	864
801	801	30	16	1.9	0.20	0.25	0.9	0	1.0	0	-0.04	801
795	795	26	16	1.6	0.18	0.24	0.9	0	0.9	0	0.21	795
673	673	26	16	1.6	-0.03	0.24	0.9	0	1.0	0	0.25	673
810	810	22	16	1.4	-0.05	0.24	0.9	0	0.9	0	0.11	810
379	379	24	16	1.5	-0.15	0.24	0.9	0	0.9	0	0.18	379
413	413	43	16	2.7	1.50	0.45	1.0	0	0.9	0	0.22	413
420	420	41	16	2.6	1.37	0.38	1.0	0	0.8	0	0.31	420
755	755	42	16	2.6	1.31	0.41	1.0	0	1.1	0	0.09	755
799	799	41	16	2.6	1.16	0.38	1.0	0	0.8	0	0.27	799
851	851	41	16	2.6	1.16	0.38	1.0	0	0.8	0	0.42	851
684	684	32	14	2.3	0.91	0.32	1.0	0	0.9	0	0.35	684
750	750	39	16	2.4	0.91	0.33	1.0	0	0.8	0	0.35	750
678	678	36	16	2.3	0.83	0.29	1.0	0	1.1	0	0.28	678
793	793	36	16	2.3	0.83	0.29	1.0	0	1.0	0	0.08	793
669	669	38	16	2.4	0.80	0.31	1.0	0	1.0	0	-0.03	669
878	878	35	16	2.2	0.75	0.28	1.0	0	1.8	1	-0.11	878
884	884	34	16	2.1	0.68	0.27	1.0	0	0.9	0	0.36	884
698	698	36	16	2.3	0.62	0.29	1.0	0	0.9	0	0.30	698
882	882	33	16	2.1	0.61	0.26	1.0	0	1.0	0	0.18	882
433	433	35	16	2.2	0.54	0.28	1.0	0	0.9	0	0.12	433
713	713	21	11	1.9	0.34	0.29	1.0	0	1.1	0	0.17	713

MEAP Math 01-07-1991 20:51:11
Table D-1 Students Measurement Report (ordered by Infit) - cont'd

Num Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
425 425	32	16	2.0	0.33	0.26	1.0	0	0.9	0	0.32	425
806 806	31	16	1.9	0.26	0.25	1.0	0	1.0	0	0.34	806
680 680	27	16	1.7	0.23	0.24	1.0	0	1.0	0	-0.05	680
671 671	30	16	1.9	0.20	0.25	1.0	0	0.9	0	0.21	671
384 384	25	16	1.6	0.12	0.24	1.0	0	0.9	0	0.21	384
672 672	25	16	1.6	0.12	0.24	1.0	0	1.0	0	-0.16	672
798 798	27	16	1.7	0.02	0.24	1.0	0	0.9	0	0.15	798
813 813	26	16	1.6	-0.03	0.24	1.0	0	0.9	0	0.18	813
372 372	21	15	1.4	-0.04	0.24	1.0	0	1.0	0	0.27	372
386 386	22	16	1.4	-0.05	0.24	1.0	0	0.9	0	0.24	386
393 393	23	15	1.5	-0.15	0.24	1.0	0	1.0	0	-0.07	393
491 491	20	14	1.4	-0.20	0.25	1.0	0	1.2	0	-0.01	491
488 488	11	9	1.2	-0.29	0.32	1.0	0	1.0	0	0.09	488
676 676	21	16	1.3	-0.31	0.24	1.0	0	1.1	0	0.21	676
389 389	14	16	0.9	-0.73	0.26	1.0	0	1.0	0	-0.05	389
371 371	12	16	0.8	-0.87	0.27	1.0	0	0.9	0	0.42	371
382 382	12	16	0.8	-0.87	0.27	1.0	0	1.0	0	-0.05	382
744 744	47	16	2.9	3.21	1.03	1.1	0	2.0	1	-0.10	744
745 745	46	16	2.9	2.47	0.73	1.1	0	1.0	0	-0.06	745
709 709	42	16	2.6	1.31	0.41	1.1	0	0.8	0	0.43	709
697 697	41	16	2.6	1.16	0.38	1.1	0	0.9	0	0.36	697
666 666	38	16	2.4	1.01	0.31	1.1	0	1.0	0	0.15	666
758 758	38	16	2.4	0.80	0.31	1.1	0	0.9	0	0.33	758
883 883	38	16	2.4	0.80	0.31	1.1	0	0.9	0	0.36	883
699 699	35	16	2.2	0.75	0.28	1.1	0	1.4	0	-0.08	699
879 879	36	16	2.3	0.62	0.29	1.1	0	1.0	0	0.50	879
873 873	33	15	2.2	0.58	0.29	1.1	0	1.0	0	0.49	873
427 427	35	16	2.2	0.54	0.28	1.1	0	1.0	0	0.28	427
689 689	35	16	2.2	0.54	0.28	1.1	0	1.1	0	0.33	689
807 807	34	16	2.1	0.47	0.27	1.1	0	1.0	0	0.16	807
432 432	31	16	1.9	0.26	0.25	1.1	0	1.1	0	-0.01	432
857 857	29	15	1.9	0.24	0.26	1.1	0	1.1	0	0.32	857
856 856	25	16	1.6	0.12	0.24	1.1	0	1.1	0	0.11	856
428 428	27	16	1.7	0.02	0.24	1.1	0	1.1	0	0.13	428
481 481	22	16	1.4	-0.05	0.24	1.1	0	1.1	0	0.18	481
366 366	25	16	1.6	-0.09	0.24	1.1	0	1.1	0	-0.00	366
485 485	20	16	1.3	-0.16	0.24	1.1	0	1.0	0	-0.01	485
376 376	22	16	1.4	-0.26	0.24	1.1	0	1.2	0	-0.29	376
367 367	16	16	1.0	-0.60	0.25	1.1	0	1.0	0	0.23	367
691 691	32	13	2.5	1.17	0.38	1.2	0	1.0	0	0.37	691
804 804	41	16	2.6	1.16	0.38	1.2	0	0.9	0	0.28	804
408 408	40	16	2.5	1.02	0.35	1.2	0	1.0	0	0.34	408
762 762	40	16	2.5	1.02	0.35	1.2	0	1.1	0	0.33	762
677 677	34	16	2.1	0.68	0.27	1.2	0	1.3	0	0.28	677
815 815	30	16	1.9	0.41	0.25	1.2	0	1.2	0	-0.02	815
811 811	29	16	1.8	0.14	0.24	1.2	0	1.2	0	0.06	811
435 435	25	14	1.8	0.12	0.26	1.2	0	1.2	0	-0.06	435
675 675	28	16	1.8	0.08	0.24	1.2	0	1.2	0	0.07	675

MEAP Math 01-07-1991 20:51:11
Table D-1 Students Measurement Report (ordered by Infit) - cont'd

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
674	674	27	16	1.7	0.02	0.24	1.2	0	1.3	1	0.02	674
436	436	20	14	1.4	0.01	0.25	1.2	0	1.3	1	0.07	436
392	392	26	16	1.6	-0.03	0.24	1.2	0	1.2	0	0.23	392
874	874	26	16	1.6	-0.03	0.24	1.2	0	1.3	1	0.15	874
695	695	22	16	1.4	-0.05	0.24	1.2	0	1.2	0	0.20	695
886	886	25	16	1.6	-0.09	0.24	1.2	0	1.3	1	0.02	886
480	480	24	16	1.5	-0.15	0.24	1.2	0	1.1	0	0.29	480
377	377	12	8	1.5	-0.32	0.32	1.2	0	1.1	0	0.34	377
492	492	15	11	1.4	-0.35	0.28	1.2	0	1.1	0	-0.14	492
670	670	20	16	1.3	-0.37	0.24	1.2	0	1.2	0	0.12	670
391	391	14	13	1.1	-0.43	0.27	1.2	0	1.2	0	0.25	391
394	394	15	16	0.9	-0.46	0.25	1.2	0	1.5	1	0.01	394
483	483	13	14	0.9	-0.67	0.27	1.2	0	1.2	0	0.02	483
397	397	43	16	2.7	1.50	0.45	1.3	0	0.9	0	0.43	397
747	747	42	16	2.6	1.42	0.41	1.3	0	1.0	0	0.67	747
885	885	41	16	2.6	1.37	0.38	1.3	0	1.2	0	0.32	885
682	682	38	16	2.4	1.01	0.31	1.3	0	1.1	0	0.44	682
402	402	39	16	2.4	0.91	0.33	1.3	0	1.1	0	0.14	402
659	659	38	16	2.4	0.80	0.31	1.3	0	1.6	1	0.11	659
465	465	22	11	2.0	0.43	0.30	1.3	0	1.3	0	0.20	465
471	471	22	11	2.0	0.43	0.30	1.3	0	1.2	0	0.00	471
814	814	30	16	1.9	0.41	0.25	1.3	0	1.2	0	0.21	814
487	487	7	5	1.4	-0.29	0.43	1.3	0	1.3	0	-0.11	487
405	405	45	16	2.8	2.24	0.60	1.4	0	0.9	0	0.20	405
748	748	41	16	2.6	1.37	0.38	1.4	0	1.2	0	0.31	748
863	863	41	16	2.6	1.16	0.38	1.4	0	1.5	0	0.13	863
757	757	44	16	2.8	1.73	0.51	1.5	0	1.4	0	-0.01	757
395	395	41	16	2.6	1.37	0.38	1.5	0	1.4	0	0.21	395
396	396	42	16	2.6	1.31	0.41	1.5	0	1.2	0	0.36	396
679	679	42	16	2.6	1.31	0.41	1.5	0	1.4	0	0.01	679
421	421	40	16	2.5	1.02	0.35	1.5	0	1.2	0	0.38	421
417	417	44	16	2.8	1.94	0.51	1.6	0	1.0	0	0.43	417
442	442	32	16	2.0	0.33	0.26	1.3	1	1.4	1	-0.04	442
441	441	28	16	1.8	0.29	0.24	1.3	1	1.3	1	0.13	441
429	429	30	16	1.9	0.20	0.25	1.3	1	1.3	1	0.35	429
388	388	29	16	1.8	0.14	0.24	1.3	1	1.3	1	0.03	388
385	385	25	16	1.6	0.12	0.24	1.3	1	1.3	1	-0.25	385
375	375	17	11	1.5	0.02	0.28	1.3	1	1.3	0	-0.29	375
387	387	23	16	1.4	0.01	0.24	1.3	1	1.3	1	-0.17	387
800	800	26	16	1.6	-0.03	0.24	1.3	1	1.3	1	0.05	800
868	868	26	16	1.6	-0.03	0.24	1.3	1	1.3	1	0.16	868
373	373	22	16	1.4	-0.05	0.24	1.3	1	1.3	1	0.02	373
390	390	20	16	1.3	-0.16	0.24	1.3	1	1.6	2	-0.21	390
431	431	23	16	1.4	-0.20	0.24	1.3	1	1.2	1	-0.11	431
881	881	31	16	1.9	0.47	0.25	1.4	1	1.4	1	0.11	881
368	368	28	16	1.8	0.08	0.24	1.4	1	1.4	1	-0.12	368
461	461	24	16	1.5	0.06	0.24	1.4	1	1.4	1	-0.03	461
403	403	40	16	2.5	1.23	0.35	1.5	1	1.4	0	0.10	403

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
712	712	40	16	2.5	1.02	0.35	1.5	1	2.1	1	0.12	712
746	746	40	16	2.5	1.02	0.35	1.5	1	1.3	0	0.34	746
850	850	38	16	2.4	1.01	0.31	1.5	1	1.4	0	0.28	850
412	412	38	16	2.4	0.80	0.31	1.5	1	1.4	0	0.28	412
409	409	35	15	2.3	0.76	0.32	1.5	1	1.3	0	0.27	409
792	792	30	15	2.0	0.34	0.27	1.5	1	1.5	1	0.04	792
866	866	41	16	2.6	1.37	0.38	1.6	1	1.3	0	0.24	866
756	756	36	16	2.3	0.62	0.29	1.6	1	1.5	1	0.32	756
855	855	32	16	2.0	0.33	0.26	1.6	1	1.5	1	0.05	855
370	370	16	16	1.0	-0.60	0.25	1.6	1	1.6	1	-0.27	370
760	760	44	16	2.8	1.73	0.51	1.7	1	1.1	0	0.38	760
406	406	41	16	2.6	1.16	0.38	1.7	1	1.3	0	0.42	406
859	859	38	16	2.4	0.80	0.31	1.7	1	1.5	0	0.27	859
854	854	43	16	2.7	1.50	0.45	1.8	1	1.6	0	0.21	854
414	414	46	16	2.9	2.47	0.73	1.9	1	1.0	0	0.23	414
398	398	43	16	2.7	1.50	0.45	1.9	1	1.4	0	0.35	398
407	407	41	16	2.6	1.37	0.38	1.9	1	1.6	1	-0.01	407
419	419	45	16	2.8	2.03	0.60	2.2	1	1.1	0	0.39	419
860	860	45	16	2.8	2.03	0.60	2.2	1	1.1	0	0.39	860
796	796	25	16	1.6	-0.09	0.24	1.5	2	1.6	2	-0.21	796
484	484	30	16	1.9	0.20	0.25	1.6	2	1.6	1	-0.17	484
426	426	25	16	1.6	-0.09	0.24	1.6	2	1.6	2	-0.23	426
374	374	25	16	1.6	0.12	0.24	1.7	2	1.7	2	0.06	374
482	482	16	14	1.1	-0.46	0.26	1.8	2	1.9	2	0.04	482

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
Count:	Mean:	29.6	14.4	2.0	0.55	0.34	1.0	-0.0	0.9	-0.0	0.21	
263	S.D.:	11.1	3.6	0.5	0.73	0.15	0.4	1.0	0.4	0.9	0.19	

RMSE 0.38 Adj S.D. 0.62 Separation 1.66 Reliability 0.73
Fixed (all same) chi-square: 878.04 d.f.: 259 significance: 0.00
Random (normal distribution) chi-square: 221.11 d.f.: 258 significance: .95

MEAP Math 01-07-1991 20:51:11
Table D-2 Items Measurement Report (ordered by infit)

Nu	Items	Score	Count	Average	Calib. Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Nu
2	2	496	227	2.2	-0.13	0.07	0.8	-2	0.8	-1	0.40	2
4	4	577	251	2.3	-0.20	0.07	0.8	-1	0.7	-1	0.36	4
7	7	568	247	2.3	-0.30	0.07	0.8	-1	0.7	-2	0.37	7
5	5	440	245	1.8	0.00	0.07	0.9	-1	0.8	-1	0.36	5
15	15	381	217	1.8	0.46	0.08	1.0	0	1.0	0	0.28	15
9	9	405	240	1.7	0.39	0.07	1.0	0	1.1	0	0.25	9
13	13	422	217	1.9	0.20	0.08	1.1	0	1.0	0	0.25	13
16	16	435	217	2.0	0.18	0.07	1.1	0	1.1	0	0.26	16
1	1	480	230	2.1	-0.12	0.07	1.0	0	1.0	0	0.26	1
14	14	478	217	2.2	-0.12	0.08	0.9	0	0.9	0	0.31	14
8	8	512	245	2.1	-0.14	0.07	1.0	0	0.9	0	0.29	8
6	6	576	256	2.3	-0.50	0.08	1.0	0	0.9	0	0.29	6
3	3	657	251	2.6	-0.71	0.08	0.9	0	0.6	-1	0.31	3
10	10	479	237	2.0	0.07	0.07	1.1	1	1.1	1	0.20	10
11	11	461	232	2.0	-0.00	0.07	1.1	1	1.2	1	0.26	11
12	12	314	217	1.4	0.92	0.08	1.6	6	1.7	6	0.00	12

Nu	Items	Score	Count	Average	Calib. Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Nu
Count:	Mean:	480.1	234.1	2.0	-0.00	0.07	1.0	0.1	1.0	0.0	0.28	
16	S.D.:	82.7	13.8	0.3	0.37	0.00	0.2	2.0	0.2	2.1	0.09	

RMSE 0.07 Adj S.D. 0.37 Separation 4.95 Reliability 0.96
Fixed (all same) chi-square: 378.57 d.f.: 15 significance: 0.00
Random (normal distribution) chi-square: 14.95 d.f.: 14 significance: .38

MEAP MATH; SHORT VERSION GRADE 10 "BIGSCALE" RASCH ANALYSIS VER. 1.73 Jan 31 20:25:51 1991
INPUT: 262 PERSONS 16 ITEMS ANALYZED: 249 PERSONS 16 ITEMS 2 CATEGORIES TABLE E-1

SUMMARY OF 249 MEASURED PERSONS

	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	7.0	14.8	-.27	.68	1.0	.0	1.1	.1
S.D.	3.8	3.0	1.33	.17	.2	.9	.8	.9

RMSE .70 ADJ.S.D. 1.13 PERSON SEP 1.62 PERSON SEP REL. .72

SUMMARY OF 16 CALIBRATED ITEMS

	COUNT	SAMPLE	CALIBRTN	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	108.4	230.1	.00	.16	1.0	-.2	1.1	.1
S.D.	43.2	11.7	1.08	.02	.2	2.1	.6	2.1

RMSE .16 ADJ.S.D. 1.07 ITEM SEP 6.49 ITEM SEP REL. .98

SUMMARY OF CALIBRATED STEPS

LABEL	VALUE	COUNT	RESIDUAL
0	0	1948	.0
1	1	1734	.0

MEAP MATH; PARTIAL CREDIT OF SHORT VERSION GRADE 10 "BIGSCALE" RASCH ANALYSIS VER. 1.73 Jan 31 20:28:18 1991
INPUT: 262 PERSONS 16 ITEMS ANALYZED: 258 PERSONS 16 ITEMS 4 CATEGORIES TABLE E-2

SUMMARY OF 258 MEASURED PERSONS

	COUNT	TEST	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	29.6	14.4	.60	.34	1.0	.1	1.0	.0
S.D.	11.0	3.6	.72	.15	.4	1.0	.4	1.0

RMSE .37 ADJ.S.D. .62 PERSON SEP 1.66 PERSON SEP REL. .73

SUMMARY OF 16 CALIBRATED ITEMS

	COUNT	SAMPLE	CALIBRTN	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	477.4	232.4	.00	.08	1.0	.3	1.0	.0
S.D.	85.1	13.8	.40	.01	.2	2.0	.2	1.9

RMSE .08 ADJ.S.D. .39 ITEM SEP 5.17 ITEM SEP REL. .96

SUMMARY OF CALIBRATED STEPS

C LABEL	CATEGORY STEP VALUE	OBSERVED COUNT	STEP CALIBR.	STEP ERROR	EXPECTED SCORE CALIBRATIONS STEP-.5	AT STEP	STEP+.5	COUNT RESID.
0	0	469	NONE			EXTREME	-1.07	-.2
1	1	595	-.16	.05	-1.07	-.46	-.01	.1
2	2	921	-.03	.04	-.01	.44	1.08	.1
3	3	1734	.19	.04	1.08	EXTREME		.0

FACETS Version No. 2.37 Copyright (c) 1990, John M. Linacre
MEAP Math 01-07-1991 20:51:11 TABLE E-3

Student Summary Report

Num	Students	Score	Count	Average	Measure Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Num
Count:	Mean:	29.6	14.4	2.0	0.55	0.34	1.0	-0.0	0.9	-0.0	0.21	
263	S.D.:	11.1	3.6	0.5	0.73	0.15	0.4	1.0	0.4	0.9	0.19	

RMSE 0.38 Adj S.D. 0.62 Separation 1.66 Reliability 0.73
Fixed (all same) chi-square: 878.04 d.f.: 259 significance: 0.00
Random (normal distribution) chi-square: 221.11 d.f.: 258 significance: .95

Item Summary Report

Nu	Items	Score	Count	Average	Calib. Logit	Model Error	Infit MnSq	Infit Std	Outfit MnSq	Outfit Std	PtBis	Nu
Count:	Mean:	480.1	234.1	2.0	-0.00	0.07	1.0	0.1	1.0	0.0	0.28	
16	S.D.:	82.7	13.8	0.3	0.37	0.00	0.2	2.0	0.2	2.1	0.09	

RMSE 0.07 Adj S.D. 0.37 Separation 4.95 Reliability 0.96
Fixed (all same) chi-square: 378.57 d.f.: 15 significance: 0.00
Random (normal distribution) chi-square: 14.95 d.f.: 14 significance: .38